



Universidad
Carlos III de Madrid

Departamento de Ingeniería de Sistemas y Automática

PROYECTO FIN DE CARRERA

DETECCIÓN Y SEGUIMIENTO DE PEATONES MEDIANTE FUSIÓN SENSORIAL

INGENIERÍA TÉCNICA INDUSTRIAL. ESPECIALIDAD EN
ELECTRÓNICA INDUSTRIAL

Autor: Jorge Guerra del Castillo

Tutor: Fernando García Fernández



Leganés, abril de 2012





Título: Detección y seguimiento de peatones mediante fusión sensorial

Autor: Jorge Guerra del Castillo

Director: Fernando García Fernández

EL TRIBUNAL

Presidente: _____

Vocal: _____

Secretario: _____

Realizado el acto de defensa y lectura del Proyecto Fin de Carrera el día ____ de _____
de 20__ en Leganés, en la Escuela Politécnica Superior de la Universidad Carlos III de
Madrid, acuerda otorgarle la CALIFICACIÓN de



VOCAL

SECRETARIO

PRESIDENTE



Agradecimientos

A todos los que me han apoyado hasta el final



Índice de contenidos

| | |
|--|----|
| Índice de contenidos | 7 |
| Índice de ilustraciones..... | 12 |
| Índice de tablas..... | 17 |
| Capítulo 1. Introducción | 18 |
| 1.1 Objetivos del proyecto..... | 18 |
| 1.2 Siniestralidad en la carretera | 18 |
| 1.2.1 Los peatones | 21 |
| 1.2.2 Mortalidad de peatones en España | 23 |
| 1.3 Evolución de los sistemas de seguridad para los automóviles..... | 25 |
| 1.4 Fusión sensorial..... | 27 |
| 1.5. Propósito final | 29 |
| Capítulo 2. Estado del arte | 30 |
| 2.1 Sistemas ADAS (Advanced Driver Assistance Systems) | 30 |
| 2.1.1 Sistema de detección de objetos en ángulo muerto | 30 |
| 2.1.2 Asistente para cambios de carril | 32 |
| 2.1.3 Sistema de visión nocturna | 33 |
| 2.1.4 Alerta por cambio involuntario de carril | 34 |
| 2.1.5 Control de crucero adaptativo (ACC) | 35 |
| 2.1.6 Sistemas de Control de la Presión de los Neumáticos..... | 37 |
| 2.1.7 Sistema de detección de señales | 39 |
| Capítulo 3. Estudio sobre la detección de peatones..... | 41 |
| 3.1 Sensor Fusion-Based Pedestrian Collision Warning System with Crosswalk Detection | 41 |
| 3.1.1 Introducción | 41 |
| 3.1.2 Características del vehículo de pruebas | 42 |
| 3.1.3 Diseño del sistema de detección de peatones y alerta..... | 43 |
| 3.1.4 Resultados experimentales | 52 |
| 3.2 Hybrid Fusion Scheme for Pedestrian Detection based on Láser Scanner and Far Infrared Camera. | 54 |
| 3.2.1 Introducción | 54 |
| 3.2.2 Algoritmo del subsistema del láser | 55 |
| 3.2.3 Detección de bajo nivel | 55 |
| 3.2.4 Clasificación de alto nivel | 58 |



| | | |
|---------|--|----|
| 3.2.5 | Subsistema de infrarrojo lejano..... | 59 |
| 3.2.5.1 | Modelo de cámara | 59 |
| 3.2.5.2 | Algoritmo de detección de peatones | 59 |
| 3.2.6 | Resultados..... | 61 |
| 3.3 | A New Approach to Urban Pedestrian Detection for Automatic Braking..... | 62 |
| 3.3.1 | Introducción | 62 |
| 3.3.2 | Foco de atención..... | 65 |
| 3.3.3 | Fusión de datos..... | 71 |
| 3.3.3.1 | Clasificación..... | 71 |
| 3.3.4 | Resultados..... | 73 |
| 3.4 | Camera and Imaging Radar Feature Level Sensorfusion for Night Vision Pedestrian Recognition..... | 78 |
| 3.4.1 | Introducción | 76 |
| 3.4.2 | Sensores | 76 |
| 3.4.3 | Sistema de fusion..... | 77 |
| 3.4.4 | Resultados..... | 80 |
| 3.5 | Pedestrian detection based on maximally stable extremal regions..... | 85 |
| 3.5.1 | Introducción | 83 |
| 3.5.2 | Sensores | 83 |
| 3.5.3 | Hipótesis sobre la presencia de peatones | 84 |
| 3.5.3.1 | Detección de áreas cálidas | 84 |
| 3.5.3.2 | Extracción de las cabezas de la imagen | 85 |
| 3.5.3.3 | Definición de las regiones de interés | 85 |
| 3.5.4 | Resultados..... | 87 |
| 3.6 | Non-parametric Laser and Video Data Fusion: Application to Pedestrian Detection in Urban Environment | 91 |
| 3.6.1 | Introducción | 91 |
| 3.6.2 | Sensores..... | 92 |
| 3.6.3 | Fusión de los datos de los sensores | 92 |
| 3.6.3.1 | Arquitectura del sistema | 92 |
| 3.6.3.2 | Asociación no-paramétrica de datos..... | 93 |
| 3.6.3.3 | Cálculo del factor de la confianza en la fusión de datos | 95 |
| 3.6.4 | Experimentos | 98 |
| 3.6.4.1 | Simulaciones..... | 98 |



| | |
|--|-----|
| 3.6.4.2 Experimentos con datos reales | 99 |
| 3.7 An application of data fusion combining laser scanner and vision in real-time driving environment recognition system..... | 105 |
| 3.7.1 Introducción | 103 |
| 3.7.2 Estructura del sistema..... | 103 |
| 3.7.3 Calibración del sistema | 104 |
| 3.7.4 Fusión de datos..... | 107 |
| 3.7.5 Resultados experimentales | 108 |
| 3.8 Comparison between GMM and KDE Data Fusion methods for Particle Filtering: Application to Pedestrian Detection from Laser and Video Measurements | 112 |
| 3.8.1 Introducción | 112 |
| 3.8.2 Información general..... | 112 |
| 3.8.3 Algoritmos de fusión de partículas..... | 114 |
| 3.8.3.1 Método GMM..... | 115 |
| 3.8.3.2 Método KDE | 116 |
| 3.8.4 Resultados..... | 118 |
| 3.9 Fast Pedestrian Detection in Dense Environment with a Laser Scanner and a Camera | 121 |
| 3.9.1 Introducción | 121 |
| 3.9.2 Principios del algoritmo | 121 |
| 3.9.3 Escáner laser..... | 122 |
| 3.9.4 Sistema de visión | 125 |
| 3.9.5 Fusión de datos | 126 |
| 3.9.6 Resultados..... | 129 |
| 3.10 Fusing Vision and Lidar – Synchronization, Correction and Occlusion Reasoning | 131 |
| 3.10.1 Introducción | 131 |
| 3.10.2 Adquisición de datos, corrección y sincronización | 131 |
| 3.10.3 Fusión de datos..... | 134 |
| 3.10.3.1 Transformación y asignación de color | 134 |
| 3.10.3.2 Tratamiento de la oclusión | 136 |
| 3.10.4 Evaluación y resultados..... | 137 |
| Capítulo 4.Descripción del sistema a desarrollar | 140 |
| 4.1 Descripción general..... | 140 |
| 4.2 Descripción del Hardware utilizado..... | 141 |
| 4.2.1 Vehículo de pruebas | 141 |



| | |
|---|-------------------------------|
| 4.2.2 Sensores..... | 142 |
| 4.2.2.1 Sensor láser | 142 |
| 4.2.2.2 Cámara de visión..... | 144 |
| 4.2.2.3 Colocación de los sensores | 145 |
| 4.2.3 Sensor GPS y sensor inercial..... | 146 |
| 4.2.4 Pantalla de visualización | 148 |
| 4.2.5 Comunicaciones | 149 |
| 4.3 Descripción del software que se va a utilizar | 149 |
| 4.3.1 Matlab | 149 |
| 4.4 Calibración de los sensores y sincronización..... | 150 |
| 4.5 Adquisición de datos | 152 |
| 4.5.1 Sistema del escáner láser | 152 |
| 4.5.2. Sistema de visión en estéreo | 157 |
| 4.5.3. Fusión de datos. | 161 |
| 4.6 Sistema de alerta | 163 |
| Capítulo 5. Resultados | 169 |
| 5.1. Introducción | 169 |
| 5.2 Secuencias analizadas mediante el filtro de Kalman | 170 |
| 5.2.1. Peatón 1..... | 173 |
| 5.2.1.1 Secuencia 1..... | 173 |
| 5.2.1.2 Secuencia 2..... | 175 |
| 5.2.2 Resto de peatones..... | ¡Error! Marcador no definido. |
| 5.3 Secuencias por Filtro de Kalman Unscented y Filtro de Partículas..... | 178 |
| 5.3.1 Peatón 1 analizado por Filtro de Kalman Unscented | 178 |
| 5.3.1.1 secuencia 1 | 178 |
| 5.3.1.2 secuencia 2 | 178 |
| 5.3.2 Peatón 1 analizado por Filtro de Partículas | 181 |
| 5.3.2.1 secuencia 1 | 181 |
| 5.3.2.2 secuencia 2 | 183 |
| Capítulo 6. Conclusiones y trabajos futuros | 185 |
| 6.1 Conclusiones..... | 185 |
| 6.2. Trabajos futuros..... | 185 |
| Capítulo 7. Presupuesto | 187 |
| 7.1 Coste de material | 187 |



| | |
|--------------------------------------|------------|
| 7.2 Costes de personal | 187 |
| 7.3. Presupuesto Total | 188 |
| Capítulo 8. Bibliografía..... | 189 |

Índice de ilustraciones

| | |
|---|----|
| Ilustración 1. Peatones cruzando paso de cebra en un entorno urbano. | 19 |
| Ilustración 2. Fallecidos en zona urbana 2001 a 2010. | 19 |
| Ilustración 3. Fallecidos en zona urbana. | 20 |
| Ilustración 4. Muertes de peatones de 2001 a 2010. | 21 |
| Ilustración 5. Número de peatones fallecidos en función de la edad. Carretera. | 22 |
| Ilustración 6. Número de peatones fallecidos por edad. Zona urbana | 22 |
| Ilustración 7. Peatones víctimas según acción del peatón. | 23 |
| Ilustración 8. Tasa de riesgo de peatones por países en Europa (2008). | 24 |
| Ilustración 9. Clasificación de los pasos de peatones por ciudades. | 25 |
| Ilustración 10. Distancia del ángulo muerto de un vehículo | 30 |
| Ilustración 11. Imagen vista por un espejo retrovisor | 31 |
| Ilustración 12. Alerta en una aproximación lenta. | 32 |
| Ilustración 13. Alerta en una aproximación rápida. | 32 |
| Ilustración 14. Visión nocturna | 33 |
| Ilustración 15. Sensores para detectar las líneas de carril | 35 |
| Ilustración 16. Control de crucero adaptativo | 36 |
| Ilustración 17. Sistema de detección de presión de neumáticos | 39 |
| Ilustración 18. Sistema de detección de señales | 40 |
| Ilustración 19. Diagrama de bloques del sistema de fusión. | 42 |
| Ilustración 20. Vehículo de pruebas utilizado y colocación en el mismo de los sensores. | 43 |
| Ilustración 21. Descripción gráfica de la relación en cruz. | 43 |
| Ilustración 22. Diagrama de flujo del algoritmo de detección de peatones con su correspondiente imagen para los bloques principales. | 45 |
| Ilustración 23. Diagrama de flujo del algoritmo de fusión de la cámara y el radar. | 46 |
| Ilustración 24. Descripción gráfica de la protección de los peatones previstos. | 48 |
| Ilustración 25. Definición de la región de interés y el nivel de riesgo | 49 |
| Ilustración 26. Información presentada en el <i>display</i> | 50 |
| Ilustración 27. Advertencia de colisión de peatones cuando la situación es crítica. | 51 |
| Ilustración 28. Descripción de las interfaces de advertencia de colisión de peatones. | 51 |
| Ilustración 29. Condiciones experimentales. | 52 |
| Ilustración 30. Resultado experimental de la detección de peatones. | 53 |
| Ilustración 31. Tránsito de la posición predicha para peatones (a vista de pájaro). | 53 |
| Ilustración 32. Diagrama de flujo de todo el sistema. | 54 |
| Ilustración 33. Procesos involucrados en el algoritmo de detección y clasificación usando el escáner láser. | 55 |
| Ilustración 34. Modelos para movimientos perpendiculares al láser. | 57 |
| Ilustración 35. Modelos para movimientos paralelos al láser. | 57 |
| Ilustración 36. Modelo general para más de tres polilíneas. | 58 |
| Ilustración 37. (a) Imagen recibida de la cámara. (b) Imagen binarizada con una cabeza. (c) Región de interés seleccionada. | 60 |
| Ilustración 38. (Primera fila) Situaciones consideradas. (Segunda fila) Áreas de interés consideradas para que la detección de un peatón sea posible. (a) Un cruce peatonal está oculto por un vehículo estacionado. (b) Un peatón está cruzando la calle detrás de un autobús detenido. (c) Un peatón aparece entre dos vehículos estacionados. (d) Un peatón está cruzando la calle entre dos vehículos detenidos en el otro lado de la carretera. Es importante tener en | |

| | |
|---|----|
| cuenta que las situaciones (a) y (b) se refieren a determinados lugares establecidos en zonas urbanas (pasos de cebra y paradas de autobús). Situaciones (c) y (d) puede ocurrir en cualquier parte de la red de carreteras. | 63 |
| Ilustración 39. Peatón parcialmente oculto por los coches aparcados que no puede ser detectado por un escáner láser colocado en el parachoques delantero, pero puede ser detectado mediante la visión, aunque esté parcialmente ocluido. | 63 |
| Ilustración 40. Detalle de la defensa delantera que muestra la integración de escáner láser y los faros. | 64 |
| Ilustración 41. Detalle de cómo la cámara está instalada dentro de la cabina..... | 64 |
| Ilustración 42. Un vehículo en movimiento y su dirección de movimiento. Las cuatro líneas paralelas corresponden a su parachoques trasero, mientras que la única línea corresponde a su lado izquierdo. Los puntos azules son los puntos de inicio de polilínea, los puntos de cian son los puntos de esquina, los puntos rojos son los puntos finales de polilínea, y los puntos de color violeta son puntos genéricos de línea. | 67 |
| Ilustración 43. Clasificación de obstáculos. Todas las clasificaciones posibles basadas en la forma están representadas en la imagen. | 68 |
| Ilustración 44. Clasificación de obstáculos. Se muestran todas las posibles clasificaciones basadas en el movimiento. Las polilíneas en movimiento cambian su posición en el sistema de coordenadas. La zona de conducción está destacada en color rosa. | 70 |
| Ilustración 45. Áreas peligrosas y posibles peatones. (Círculos) Áreas peligrosas. (Cuadrados) Posibles peatones. | 71 |
| Ilustración 46. Dos <i>frames</i> de una secuencia. (a) El sistema detecta el peatón en la parte ocluida y emite una alerta interna, pero sin previo aviso al conductor. (b) Cuando el peatón detectado es totalmente visible sigue la alerta interna, y se envía un mensaje de advertencia de inmediato al conductor. | 73 |
| Ilustración 47. Se muestran algunos peatones que aparecieron de repente y fueron detectados correctamente (a) en un aparcamiento subterráneo, (b) bajo la lluvia, (c) detrás de un vehículo mal alineado, (d) tras una pared, y (e) en la noche, y (f) a un peatón que apareció de repente detectado como un peatón no peligroso (falso negativo)..... | 75 |
| Ilustración 48. (a) Investigación de vehículos con sensor de posiciones de montaje y (b) generación avanzada del sensor radar ACC ARS300 de Continental AG. | 76 |
| Ilustración 49. Clasificador en cascada: Cada etapa descarta hipótesis y pasa a la siguiente llevando muestras a la etapa sucesivas y más complejas..... | 78 |
| Ilustración 50. Las mediciones de radar: la imagen de la izquierda representa la escena tomada por la cámara. La imagen centrada muestra la intensidad de un umbral de radar (valor) y la velocidad relativa (tono). Los tres peatones en movimiento pueden ser reconocidos por el componente verde del color alrededor de 41m, 75m y 108m. La imagen derecha muestra todos los valores de intensidad de radar (el azul corresponde baja intensidad y el rojo a una mayor intensidad). | 79 |
| Ilustración 51. Análisis de costo computacional en las primeras etapas de la cascada hasta la capa 12: El detector de imagen de referencia basado en la fusión y el detector clasifican el mismo conjunto de hipótesis. El detector de fusión muestra resultados también similares en ambientes altamente estructurados tales como las zonas suburbanas, con edificios y en gran medida los cambios de condiciones de luminosidad..... | 80 |
| Ilustración 52. Comparación de rendimiento entre el detector basado en la imagen de referencia detector (línea continua) y el detector de fusión (línea discontinua) para dos coberturas 0,5 (azul, los dos gráficos superiores) y 0,6 (cian, gráficos inferiores). | 81 |
| Ilustración 53. Se pueden ver ejemplos de detección en un entorno suburbano (imagen de la cámara con detecciones verde a la izquierda y los datos de radar con detecciones cian derecha): (a) Típico resultado del detector basado en la imagen de referencia. La incertidumbre en el espacio de localización es muy alta con características sólo para la clasificación de la imagen. Las | |

| | |
|--|-----|
| figuras (b) y (c) muestran el mayor rendimiento del detector de fusión que incorpora también las características proporcionadas por el sistema de radar. El resultado es una localización precisa de los peatones en (b) 65m y 100m (c). | 82 |
| Ilustración 54. Configuración de la plataforma del sensor..... | 84 |
| Ilustración 55. Detección de áreas cálidas. (a) Imagen original. (b) Curvas de nivel sobre la base de la MSER. (c) Formas rectangulares, como consecuencia del enfoque MSER. | 85 |
| Ilustración 56. Valores típicos de la medida de dispersión y desorden. | 86 |
| Ilustración 57. Proceso de los resultados de MSER | 87 |
| Ilustración 58. Extracción de MSER en base al umbral: (a) número de regiones extraídas, (b) tasa de regiones extraídas que contienen peatones. | 88 |
| Ilustración 59. Resultados de la detección: (d) muestra un falso positivo. | 89 |
| Ilustración 60. Curvas ROC para los resultados basados en (a) medida de dispersión y (b) plantilla de unión. | 90 |
| Ilustración 61. Localización de los sensores en el vehículo de prueba | 92 |
| Ilustración 62. Arquitectura multi-módulo con información de lidar y visión para la detección de peatones y su clasificación. Destacan en color rojo las principales aportaciones. | 93 |
| Ilustración 63. Un ejemplo de una mezcla de distribuciones de Gauss y uniforme, en azul la distribución uniforme, de color verde la distribución de Gauss y en rojo la mezcla de las distribuciones. A la izquierda, $s = 0,15$ m y $DOR = 0,95$. A la derecha, $s = 0,15$ m y $DOR = 0,55$ | 97 |
| Ilustración 64. Asociación de datos de varios sensores. Aquí, de acuerdo con el criterio del vecino más cercano, O1 se asociaría a l objeto 1 y O3 al objeto 2, la asociación correcta está dada por el algoritmo de Parzen..... | 99 |
| Ilustración 65. Ejemplo de cálculo de probabilidad en una nube de partículas en la ilustración 33. En la izquierda, $s = 0,15$ m y $DOR = 0,95$ y en la derecha, $s = 0,15$ m y $DOR = 0,55$ | 99 |
| Ilustración 66. Ejemplo de detección en una sección transversal después de una fusión centralizada de datos de imágenes lidar y de vídeo. El punto rojo representa la detección lidar y los rectángulos azules representan cámara de detección. Los rectángulos amarillos son los resultados proporcionados por el módulo de fusión de datos. | 101 |
| Ilustración 67. Ejemplo de detección en un aparcamiento después de una fusión centralizada de datos de imágenes lidar y de vídeo. Se puede observar la correcta detección de un peatón a una distancia de hasta 25 metros. | 101 |
| Ilustración 68. Ejemplo de detección en un aparcamiento después de una fusión centralizada de datos de imágenes lidar y de vídeo. Se puede observar la correcta detección de varios peatones en direcciones distintas. | 102 |
| Ilustración 69. Descripción de la estructura del sistema y la posición de sensor. 1 es la cámara y el 2 es el escáner láser. α es el grado de inclinación del escáner láser. H es la altura de montaje del escáner láser. | 104 |
| Ilustración 70. Modelo geométrico del escáner láser para la calibración extrínseca. | 105 |
| Ilustración 71. Punto de extracción de características, la relación de los tres sistemas de coordenadas. | 107 |
| Ilustración 72. La plataforma del sistema..... | 108 |
| Ilustración 73. Puntos de extracción de características y proceso de calibración. | 109 |
| Ilustración 74. Resultados de la fusión de datos..... | 111 |
| Ilustración 75. Arquitectura de la fusión de datos usando la información proporcionada por el lidar y la cámara de visión. | 113 |
| Ilustración 76. A la izquierda, el escáner láser Ibeo ALASCA XT y la cámara Marlin 146-C. A la derecha, el vehículo de pruebas IGN. | 114 |
| Ilustración 77. Ejemplo de detección, en una sección transversal después de una fusión centralizada de datos del lidar y de la imagen. Se detectan peatones en diferentes orientaciones. | 118 |

| | |
|--|-----|
| Ilustración 78. Comparativa del error medio cuadrático en posición radial entre los métodos de GMM y KDE, cuando va aumentando el número de partículas. | 119 |
| Ilustración 79. Comparación del tiempo de cálculo entre los métodos GMM y KDE (en el caso de dos sensores), cuando el número de partículas va aumentando | 119 |
| Ilustración 80. Comparativa del error medio cuadrático en posición radial entre los métodos de GMM y KDE en el seguimiento de un peatón presentado en la figura. 77..... | 120 |
| Ilustración 81. Para decidir si el objeto i se ocluye o no, se considera la posición de los dos obstáculos que le rodean (objeto $i-1$ y objeto $i+1$)...... | 122 |
| Ilustración 82 En esta figura se han calculado las señales de detección, reconocimiento y seguimiento del algoritmo del escáner láser..... | 129 |
| Ilustración 83 Aquí solamente se calcula la señal de seguimiento, en el algoritmo de visión. | 129 |
| Ilustración 84. Finalmente, mediante la fusión de ambos se pueden detectar, reconocer y seguir eficazmente los peatones de la escena. | 130 |
| Ilustración 85. . La imagen en color de la cámara (a) y la correspondiente nube de puntos corregida de 105 puntos de datos del lidar Velodyne (b). | 132 |
| Ilustración 86. Sincronización de la cámara y el lidar. La cámara tiene que ser disparada de tal manera que capture la imagen cuando los rayos láser pasen el campo de visión de la cámara (azul). En este caso se muestra cuando el vehículo está parado..... | 133 |
| Ilustración 87. Sincronización de la cámara y el lidar. La cámara tiene que ser disparada de tal manera que capture la imagen cuando los rayos láser pasen el campo de visión de la cámara (azul). En este caso se muestra cuando el vehículo está en movimiento..... | 134 |
| Ilustración 88. Algunos rayos láser alcanzan objetos en posiciones que se ocluyen a la cámara | 135 |
| Ilustración 89. Detección de la oclusión: por cada punto que se itera a través de la profundidad de lista ordenada de los cascos convexos H3. Identificación del segmento del primer casco que incluye un punto (en este caso H2) se compara con el punto de Identificación del segmento. | 137 |
| Ilustración 90. Escena observada por ambos sensores. | 138 |
| Ilustración 91. El número de píxeles de la señal mal coloreados con y sin detección de la oclusión ... | 138 |
| Ilustración 92. En esta imagen se puede ver cómo un peatón ocluido por un autobús intenta cruzar la calle..... | 140 |
| Ilustración 93. Vista lateral del vehículo de pruebas IVVI 2.0. | 141 |
| Ilustración 94. Telémetro láser SICK LMS 291 | 142 |
| Ilustración 95. Vista interna del SICK LMS 291. | 143 |
| Ilustración 96. Comportamiento del láser para un escaneo de 0,25° de resolución. | 144 |
| Ilustración 97. Cámara estéreo Bumblebee2..... | 144 |
| Ilustración 98. Imagen obtenida por la cámara BUmblebee2. | 145 |
| Ilustración 99. Posición de los sensores en el vehículo de pruebas IVVI 2.0. | 146 |
| Ilustración 100. Sensor GPS e inercial MTI-G de Xsens | 147 |
| Ilustración 101. Pantalla Xernac 705 YV de 7" | 148 |
| Ilustración 102. Posición de la pantalla en el interior del vehículo. | 148 |
| Ilustración 103. Ordenadores de a bordo del vehículo IVVI 2.0. | 149 |
| Ilustración 104. Logotipo de MATLAB | 150 |
| Ilustración 105. Sistemas de coordenadas de los sensores en el vehículo. 1 es la cámara y 2 el escáner láser, donde H es la altura y α el grado de inclinación del láser. | 151 |
| Ilustración 106. Secuencia de un peatón caminando perpendicularmente al láser. | 154 |
| Ilustración 107. Secuencia de un peatón acercándose al láser..... | 154 |
| Ilustración 108. Diagrama de flujo del algoritmo de clasificación de peatones. | 156 |
| Ilustración 109. Ejemplo de imágenes típicas que pueden ser obtenidos por la cámara, la primera fila muestra peatones, la segunda fila imágenes en las que no hay peatones y la tercera fila peatones parcialmente ocluidos. | 158 |
| Ilustración 110. Diagrama de flujo para algoritmo del subsistema de visión. | 160 |



| | |
|---|------------|
| Ilustración 111. Diagrama de flujo del algoritmo de fusión. | 162 |
| Ilustración 112. Diagrama de flujo del algoritmo de alerta. | 165 |
| Ilustración 113. Información presentada en la pantalla | 166 |
| Ilustración 114. Descripción gráfica del sistema de alerta y frenada de emergencia. | 168 |
| Ilustración 115. Detecciones totales | 170 |
| Ilustración 116. Detecciones actuales | 171 |
| Ilustración 117. Lo que saldría por pantalla | 171 |
| Ilustración 118. Histórico de las detecciones..... | 172 |
| Ilustración 119. Frame de la secuencia..... | 173 |



Índice de tablas

| | |
|---|-----|
| Tabla 1. Resultados | 74 |
| Tabla 2. Características del detector. | 80 |
| Tabla 3. Tasa de detecciones falsas y correctas: cuando sólo se utiliza la cámara o el lidar, y después de la fusión de datos. | 100 |
| Tabla 4. Resultados con los parámetros de la cámara | 110 |
| Tabla 6. Secuencia 1..... | 173 |
| Tabla 7. Secuencia 2..... | 175 |
| Tabla 8. Resto de secuencias analizadas | 176 |
| Tabla 9. Secuencia 1UKF | 179 |
| Tabla 10. Secuencia 2 UKF..... | 180 |
| Tabla 11. Secuencia 1PF | 181 |
| Tabla 12. Secuencia 2 PF | 183 |
| Tabla 13. Costes de material | 187 |
| Tabla 14. Costes de personal..... | 188 |
| Tabla 15. Coste total del Proyecto | 188 |

Capítulo 1. Introducción

1.1 Objetivos del proyecto

El presente proyecto tiene como fin el estudio de los distintos sistemas de detección y seguimiento de peatones que existen actualmente y el diseño de un sistema de detección y seguimiento de la trayectoria de peatones en las vías públicas, a partir de los datos proporcionados por un escáner láser y una cámara de visión estéreo, para su posterior procesamiento y ejecución de acciones tales como seguimiento y predicción de la trayectoria de los peatones y avisar al conductor de una posible colisión si fuera necesario. Para que la futura implementación fuera posible en su totalidad se tendría que modificar el vehículo de pruebas para ser aceptado por la legislación de tráfico. Este sistema además servirá como base para diversas aplicaciones de seguridad, conducción autónoma, etc... que necesiten conocer la posición y futura situación de los peatones que le rodean. A continuación se detallan datos que ponen en relieve la necesidad de este tipo de sistemas.

1.2 Siniestralidad en la carretera

Según un informe de la DGT en 2010 [1] los atropellos son la primera causa de muerte por accidente de tráfico en la ciudad. De forma invariable desde el 2003, los atropellos causan más del 40% de los muertos por accidente de tráfico en zona urbana y el 15% de los lesionados. Sabemos que a partir de 70 km/h un atropello es habitualmente una muerte segura, mientras que a 50 km/h el riesgo de muerte se reduce al 75% y a 30 km/h se pueden evitar tres de cada cuatro atropellos.



Ilustración 1. Peatones cruzando paso de cebra en un entorno urbano.

En el año 2010 se produjeron en España 85.503 accidentes de circulación con víctimas, entendiéndose como tal aquél en el que una o varias personas resultan muertas o heridas y está implicado al menos un vehículo en movimiento.

En esos accidentes fallecieron 2.478 personas dentro de los 30 días siguientes al accidente y 120.345 resultaron heridas, de las cuales, 11.995 lo fueron gravemente, es decir, necesitaron más de 24 horas de hospitalización.

El índice de gravedad de los accidentes en el año 2010 fue de 2,9 muertos por cada 100 accidentes con víctimas y el de letalidad de 2,02 muertos por cada 100 víctimas. En el año 2010, en el 7,8% de los accidentes mortales hubo más de un muerto y en el 0,4% de ellos hubo más de tres.

| Fallecidos en Zona Urbana | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | Distribución % 2010 | 2010/ 2009 | 2010/ 2003 | 2010/ 2001 |
|----------------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------------------|------------|-------------|-------------|
| Salida de la vía | 164 | 148 | 146 | 158 | 123 | 108 | 87 | 87 | 69 | 62 | 11% | -10% | -58% | -62% |
| Colisión frontal | 61 | 58 | 52 | 51 | 45 | 40 | 29 | 29 | 28 | 15 | 3% | -46% | -71% | -75% |
| Colisión lateral y frontolateral | 215 | 213 | 225 | 228 | 183 | 163 | 149 | 152 | 100 | 87 | 16% | -13% | -61% | -60% |
| Colisión trasera y múltiple | 50 | 51 | 40 | 37 | 42 | 38 | 63 | 38 | 44 | 30 | 5% | -32% | -25% | -40% |
| Atropello a peatón | 372 | 341 | 360 | 333 | 326 | 281 | 286 | 240 | 268 | 270 | 49% | 1% | -25% | -27% |
| Vuelco | 14 | 18 | 11 | 20 | 20 | 19 | 22 | 21 | 15 | 18 | 3% | 20% | 64% | 29% |
| Otro tipo de accidente | 98 | 83 | 85 | 73 | 51 | 88 | 105 | 67 | 60 | 68 | 12% | 13% | -20% | -31% |
| Total | 974 | 912 | 919 | 900 | 790 | 737 | 741 | 634 | 584 | 550 | 100% | -6% | -40% | -44% |

Ilustración 2. Fallecidos en zona urbana 2001 a 2010.

En lo que más nos interesa para el presente proyecto, la zona urbana, en 2010, el 49% de las víctimas mortales se produjeron por atropello, el 16% en colisiones laterales y frontolaterales y el 11% en salidas de la vía. Respecto de 2009 se han reducido considerablemente las cifras de fallecidos por colisión frontal (-46%). Por el contrario, han aumentado, los fallecidos en vuelcos (3%) y los atropellos (1%). Respecto de 2003, las reducciones más importantes se han producido en colisiones frontales y colisiones laterales y frontolaterales. Con respecto a 2001, además de las anteriores tipologías, habría que añadir las salidas de vía.

En zona urbana se registraron un total de 46.329 accidentes con víctimas, en los cuales fallecieron 550 personas (el 22% del total). Dentro de la zona urbana se incluyen las travesías, vías en las que sucedieron el 1,3% de los accidentes ocurridos en zona urbana y el 11% de las víctimas mortales. En el periodo 2001-2010, las víctimas mortales ocurridas en zona urbana han descendido un 44%, porcentaje de variación muy parecido al registrado entre 2003 y 2010 (-40%). En comparación con el año 2009 el número de víctimas mortales se ha reducido en 34 personas, lo que supone una disminución del 6%.

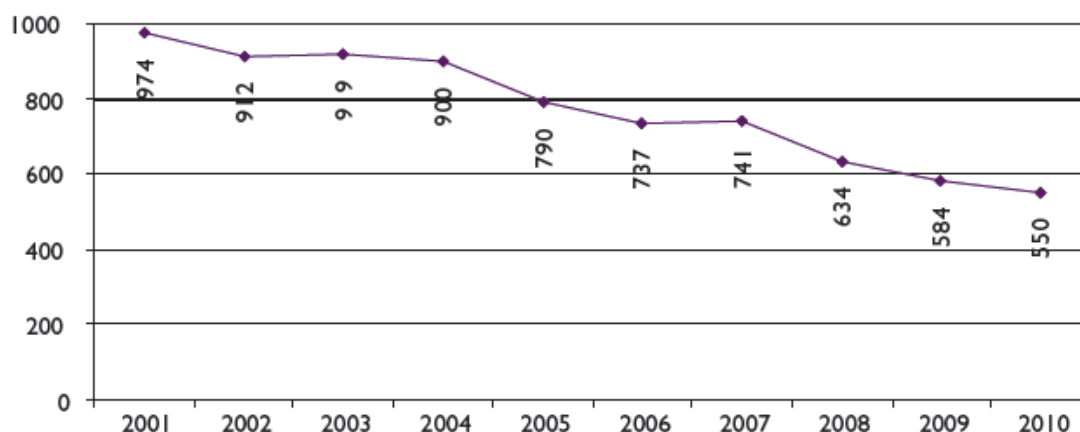


Ilustración 3. Fallecidos en zona urbana.

1.2.1 Los peatones

En 2010 fallecieron 471 peatones en accidentes de tráfico, lo que supone el 19% del total de fallecidos. Este porcentaje es superado con creces si nos situamos en el contexto de la zona urbana, donde uno de cada dos fallecidos por accidente de tráfico era un peatón. Por su parte, en carretera supone el 10% de los fallecidos totales.

| | Carretera | | Zona Urbana | | Total | |
|------|------------------|-----------------------------------|------------------|-----------------------------------|------------------|-----------------------------------|
| | Peatones muertos | % Peatones muertos sobre el total | Peatones muertos | % Peatones muertos sobre el total | Peatones muertos | % Peatones muertos sobre el total |
| 2000 | 451 | 10% | 447 | 42% | 898 | 16% |
| 2001 | 469 | 10% | 377 | 39% | 846 | 15% |
| 2002 | 433 | 10% | 343 | 38% | 776 | 15% |
| 2003 | 424 | 9% | 363 | 40% | 787 | 15% |
| 2004 | 340 | 9% | 343 | 38% | 683 | 14% |
| 2005 | 348 | 10% | 332 | 42% | 680 | 15% |
| 2006 | 317 | 9% | 296 | 40% | 613 | 15% |
| 2007 | 287 | 9% | 304 | 41% | 591 | 15% |
| 2008 | 236 | 10% | 266 | 42% | 502 | 16% |
| 2009 | 201 | 9% | 269 | 46% | 470 | 17% |
| 2010 | 193 | 10% | 278 | 51% | 471 | 19% |

Ilustración 4. Muertes de peatones de 2001 a 2010.

En el período 2001-2010 el número de peatones fallecidos en carretera se ha reducido en todos los grupos de edad con descensos superiores al 50% salvo los grupos de 45 a 54 años y de más de 74 años. El mayor descenso en porcentaje se produce en la franja de 15 a 24 años, aunque en 2010 ha habido un muerto más que en 2009.

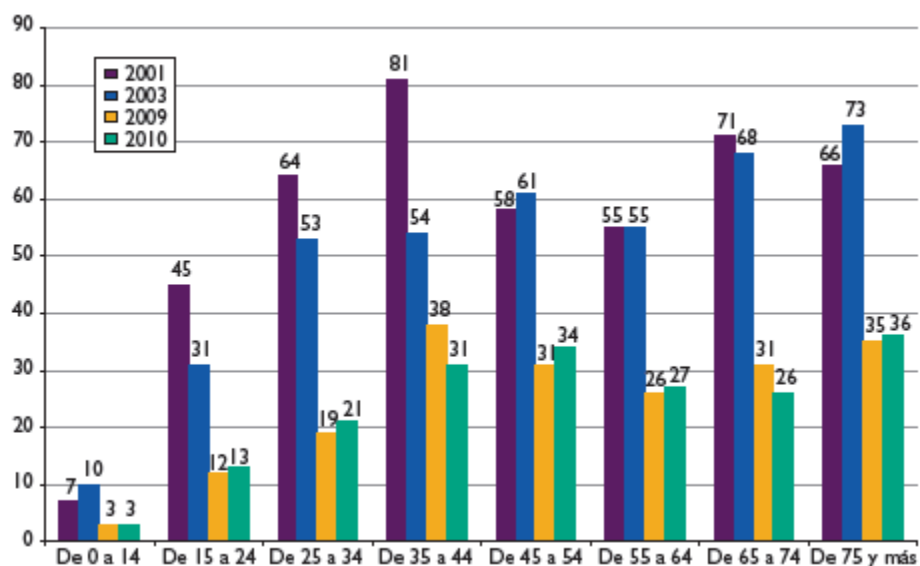


Ilustración 5. Número de peatones fallecidos en función de la edad. Carretera.

En zona urbana, el 38% del total de peatones fallecidos tenían más de 74 años (106 de los 278 peatones fallecidos). En el período 2001-2010 el mayor descenso en el número de peatones fallecidos se registra en el grupo de edad de 65 a 74 años (39%) seguido de los de 15 a 24 años y de los de 55 a 64 (37%). Con respecto a 2009, todos los grupos de edad presentan incrementos salvo el intervalo de 45 a 64.

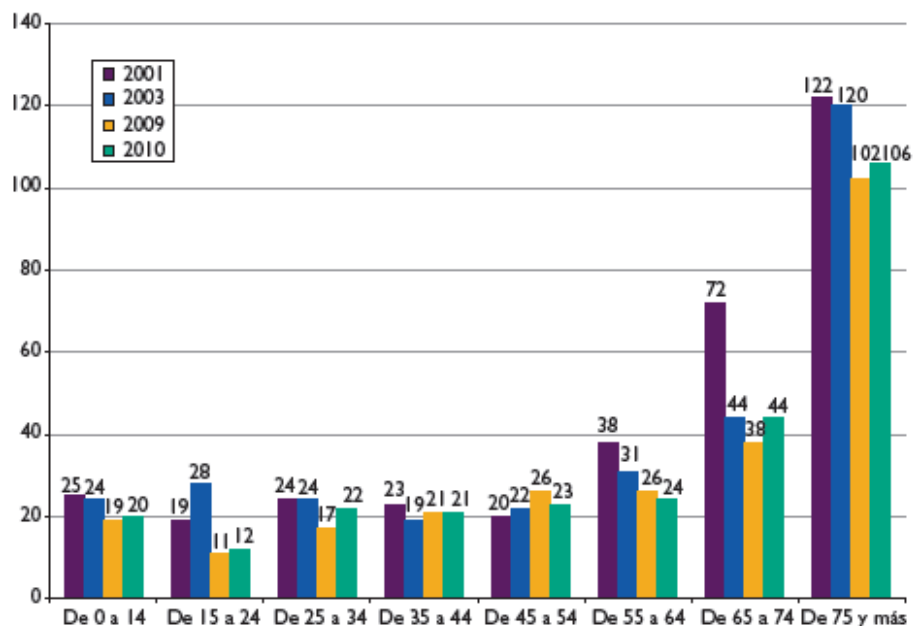


Ilustración 6. Número de peatones fallecidos por edad. Zona urbana

El 35,3% de los peatones lesionados o fallecidos en un accidente de tráfico se encontraban cruzando fuera de una intersección, mientras que un 31,6% lo hacía dentro de ella.

| | |
|-------------------------------------|-------|
| Atravesando intersección | 31,6% |
| Cruzando calzada fuera intersección | 35,3% |
| En la calzada por su derecha | 4,0% |
| En la calzada por su izquierda | 1,8% |
| Sobre acera o refugio | 5,1% |
| Otra acción | 22,3% |

Fuente: Dirección General de Tráfico (DGT).

Ilustración 7. Peadones víctimas según acción del peatón.

El 26% de las víctimas por atropello de peatones cometían alguna infracción cuando se produjo el accidente, ascendiendo hasta el 53% en el caso de los fallecidos. Por grupos de edad, el 36% de los menores de 15 años cometieron alguna infracción en el momento del accidente, mientras que para el resto de grupos el porcentaje se sitúa en el entorno del 24%.

1.2.2 Mortalidad de peatones en España

[2] Aunque España se encuentra en la franja media de países europeos en cuanto a la tasa de riesgo (peatones muertos por millón de habitantes) con 11,1 muertos por los 15,5 de media europea, la probabilidad de que un peatón muera en España es el doble que en Holanda o Suecia, países referentes en materia de seguridad vial en Europa. España se debe marcar un objetivo de reducción de la mortalidad de peatones para los próximos años que nos permita pasar de estar en una posición intermedia en el ranking europeo a ser uno de los países a la cabecera de Europa.

Se da la circunstancia que un 37,4% de las causas de los accidentes con peatones muertos en España se debe a que cruza fuera de una intersección (paso regulado).

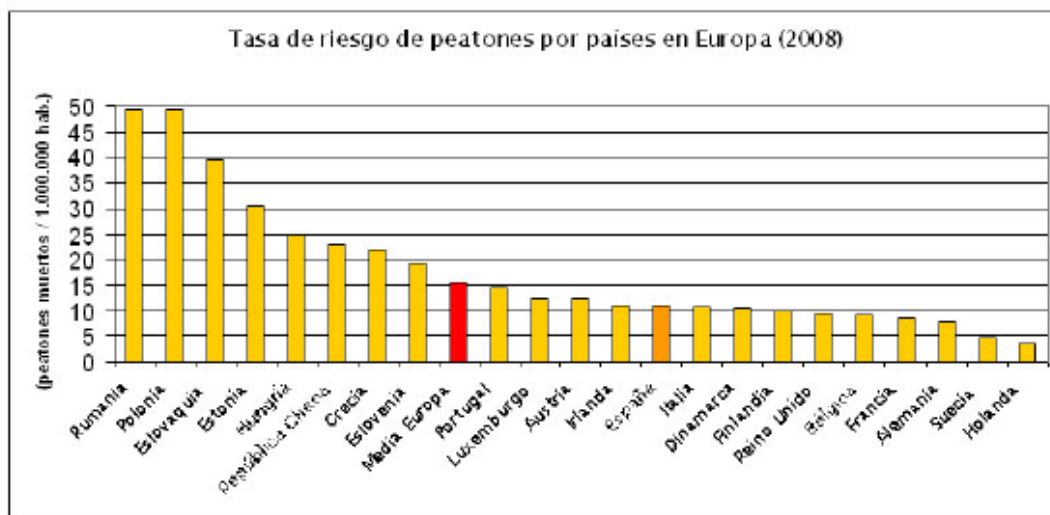


Ilustración 8. Tasa de riesgo de peatones por países en Europa (2008).

Se han analizado dos ciudades españolas, Málaga y Valencia, que ocuparon respectivamente el penúltimo y antepenúltimo lugar de la tabla de resultados de toda Europa. En Málaga, 2 pasos fueron valorados como “buenos” y 4 como “malos”; mientras que en Valencia 4 pasos fueron considerados “buenos” y 4 “malos”. A pesar de observar algunas soluciones excelentes para la accesibilidad (por ejemplo: dispositivos de cuenta atrás en los semáforos de Málaga), las condiciones de mantenimiento, el aparcamiento ilegal y la mala luz artificial contribuyeron a los resultados inadecuados obtenidos.

Además de Málaga y Valencia, se auditaron pasos de peatones en Madrid, Barcelona y Sevilla. También en la comparativa nacional estas dos ciudades obtienen resultados inferiores, lo que se traduce en la necesidad de las ciudades medianas de mejorar en los diseños de sus pasos. Sevilla saca la mejor puntuación con el único paso considerado muy bueno.

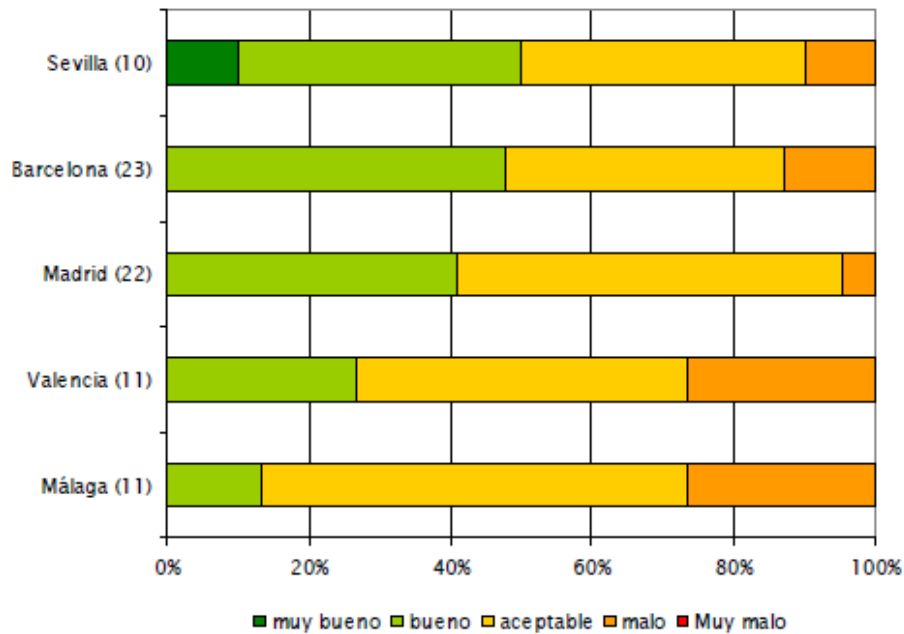


Ilustración 9. Clasificación de los pasos de peatones por ciudades.

1.3 Evolución de los sistemas de seguridad para los automóviles

Este proyecto parte de una preocupación real de la sociedad, y de su demanda de una mayor seguridad en los automóviles que propicie una reducción en el número de víctimas por accidentes como se ha visto en los estudios anteriores.

Los datos de estos estudios nos dan una visión de por qué actualmente siguen surgiendo una amplia gama de sistemas de ayuda a la conducción, independientes o cooperativos basados en tecnologías de información y comunicación, denominados ADAS (sistemas avanzados de ayuda a la conducción).

Algunos de estos sistemas están ya en el mercado, como veremos en el siguiente capítulo y otros están aún en desarrollo, todos ellos destinados a obtener una menor siniestralidad y mayor seguridad vial.

En un principio las investigaciones se centraban en la seguridad pasiva, la cual se encarga de minimizar los posibles daños de los ocupantes del vehículo en el caso de que llegue a producirse un accidente, gracias a estas investigaciones se produjeron el diseño y desarrollo de estructuras de deformación del vehículo para la mejor absorción de la energía en el caso de impacto, la incorporación de los cinturones de seguridad a los automóviles en 1950, el mejor sistema de seguridad pasiva jamás inventado (el primer vehículo que los traía de serie fue el Volvo Amazon de 1959), o el airbag en 1971, dispositivos que necesitaban un menor uso de la electrónica, todavía poco desarrollada entre las décadas de los 50 y 70.

A partir de la década de los 70 y gracias a la introducción de la informática y electrónica, tanto analógica como digital, surgen una serie de sistemas más orientados a la seguridad activa, mecanismos o dispositivos destinados a disminuir el riesgo de que se produzca un accidente. Un ejemplo son los siguientes sistemas: el ABS (Bosch 1970) cuya función es evitar que se bloqueen las ruedas al producirse un frenazo intenso, modulando automáticamente la presión sobre los frenos cuando el conductor pisa el pedal a fondo. El ESP (Bosch 1995) son las iniciales en alemán del “control electrónico de estabilidad” que detecta la desviación de la trayectoria del vehículo con respecto a la dirección que se desea llevar, el sistema frena ligeramente cada rueda por separado y así devuelve el vehículo a la trayectoria deseada. Actualmente la mayor parte de los nuevos vehículos incorporan lo que se denomina Sistema de control de tracción que garantiza la estabilidad durante la conducción, recuperando la adherencia entre neumático y asfalto cuando el conductor se excede en la aceleración. Es un sistema desarrollado sobre la base del ABS y comúnmente se denomina tanto con las siglas ASR (*Automatic Stability Control*) como TCS (*Traction Control System*).

Según un estudio de la revista “Auto Bild” con la colaboración de Ford, si todos los vehículos incorporaran ESP y ABS de serie se podría contribuir a reducir un 30% los accidentes relacionados con frenadas de emergencia y evasión de objetos en la trayectoria del vehículo.

Son tantos los beneficios de este sistema en la reducción de siniestros, que los vehículos de turismo e industriales ligeros, ya incorporan como elemento de serie el ESP a partir de noviembre de 2011 para los estados de la Unión Europea. Este dato nos puede dar una idea de la importancia de los sistemas tecnológicos incorporados en los automóviles.

Posteriormente a estos dispositivos, surgen multitud de sistemas ADAS que si bien no son tan efectivos en la reducción de víctimas como los presentados anteriormente, sirven de una gran ayuda al conductor para evitar accidentes. Estos sistemas intentan de manera individual disminuir los tipos de accidentes mortales y no mortales más frecuentes, avisando al conductor o tomando medidas preventivas siendo, según las estadísticas de la DGT para el año 2009, la salida de la vía fue el más frecuente, seguido de colisiones frontal y fronto-lateral y atropellos a peatones, por este orden.

Esto significa que un sistema como el diseñado para el reconocimiento y seguimiento de peatones se podría utilizar para evitar el cuarto tipo de accidente más habitual, incluso podrían verse reducidos otro tipo de accidentes debido a que algunos se producen por la acción de evitar el atropello a peatones, y de esta forma, mejorar la reacción de los conductores frente a posibles distracciones.

1.4 Fusión sensorial

Los sistemas de fusión sensorial son una de las principales fuentes de investigación en la actualidad, en el campo de los sistemas inteligentes de transporte. La razón de su importancia radica en la falta de sensores fiables que cumplan las elevadas exigencias de este tipo de sistemas. De esta forma, combinando varios sistemas, es posible superar las limitaciones inherentes a cada sistema por separado.

Típicamente los sensores empleados son los telémetros láser (LIDARS) y sistemas de visión artificial en cualquiera de sus diferentes configuraciones. Las razones de elegir estos sistemas de sensores es que se complementan:

El telémetro láser es capaz de medir las distancias a los objetos cercanos con una alta fiabilidad en las medidas, aunque la información que provee es limitada. Se ha demostrado que el procesamiento de los datos del laser hace que el seguimiento sea mucho más rápido y más robusto que el basado en visión, en condiciones complejas con estados de tiempo y condiciones de luz no favorables. Por el contrario el laser nos proporciona una información con unas características más pobres y con menos rasgos sobre la escena observada.

Los sistemas de visión artificial, son una fuente de información, aunque la elevada cantidad de datos que provee es el motivo de que su procesamiento sea más tedioso y complejo, y por lo tanto, su fiabilidad sea menor.

De esta forma, combinando ambos sistemas, es posible obtener un sistema sensorial que sea altamente fiable.

Los sistemas de ayuda a la conducción que emplean fusión sensorial se pueden dividir principalmente en tres tipos. Dependiendo del nivel de abstracción en el que se realiza la fusión sensorial.

Nivel bajo. Consiste en la mezcla de varias fuentes de información sin procesar, para formar un nuevo conjunto de datos que provean una información más completa [3].

Nivel medio. Consiste en extraer ciertas características de cada uno de los objetos encontrados mediante cada fuente de información por separado, como texturas, esquinas, líneas... Estas características son combinadas en fases superiores para dar una estimación de los tipos de objetos encontrados [4].

Nivel alto. Consiste en realizar una detección diferente para cada uno de los sistemas de adquisición que se están empleando, para finalmente combinar los resultados tomando una decisión del objeto más probable [5].

1.5. Propósito final

El propósito del presente proyecto es realizar un estudio detallado de los distintos sistemas de detección y seguimiento de peatones que existen mediante fusión sensorial y seleccionar las mejores características de ellos para diseñar un nuevo sistema que posiblemente pueda ser implementado en un futuro, todo ello pensado para disminuir el número de atropellos que ocurren principalmente en las zonas urbanas.

Para ello hemos seleccionado el escáner láser, la cámara de visión y el resto de sensores necesarios que dispone el vehículo de pruebas y hemos diseñado un algoritmo que permita la detección y el seguimiento de peatones en un entorno urbano mediante la fusión de los datos del láser y la cámara estéreo y además se ha diseñado un sistema de alerta con varios niveles de acuerdo a la peligrosidad y el riesgo de colisión. Como punto extra se ha añadido que el vehículo pudiese realizar una frenada de emergencia, pero para ello sería necesario realizar una modificación aprobada por la regulación de tráfico, cosa que actualmente no es posible.

Capítulo 2. Estado del arte

2.1 Sistemas ADAS (Advanced Driver Assistance Systems)

Las primeras generaciones de sistemas ADAS comenzaron a ser introducidas en el mercado desde hace relativamente muy pocos años, especialmente en vehículos de alta gama. Así, en el año 1998, Mercedes introdujo por primera vez en Europa el sistema ACC (Adaptative Cruise Control). Estos sistemas permiten al conductor realizar una conducción más confortable y al mismo tiempo avisarle si hay algún peligro y disminuir las posibilidades de que se produzca algún accidente de circulación. Como veremos en el siguiente capítulo, este proyecto se centra principalmente en uno de ellos, el sistema de detección y seguimiento de peatones.

A continuación se van a describir algunos de los sistemas ADAS utilizados actualmente en los vehículos:

2.1.1 Sistema de detección de objetos en ángulo muerto

[6] El sistema de detección de objetos en ángulo muerto tiene como objetivo alertar al conductor si hay otro vehículo en el ángulo muerto de los retrovisores o cerca de él.

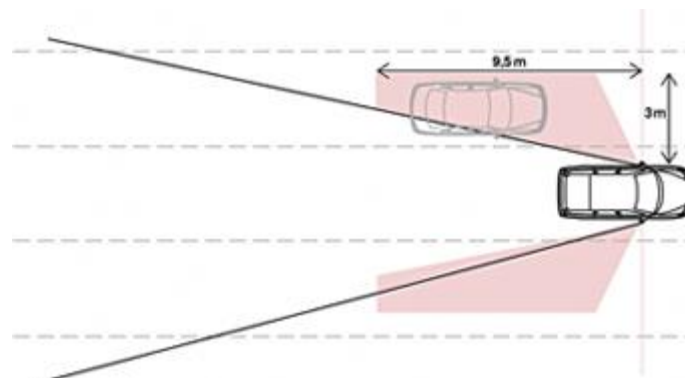


Ilustración 10. Distancia del ángulo muerto de un vehículo

Para ello, consta de sensores que vigilan constantemente la zona lateral próxima al coche. Generalmente bien son sensores radar de corto o medio alcance a 24 GHz, o bien sistemas de procesamiento de imágenes. Estos sensores proporcionan información a una centralita de control que, en caso necesario, emite un aviso acústico, visual o táctil (o una combinación de varios).

Existen sistemas que pueden alertar de forma continua de la existencia de vehículos en el ángulo muerto independientemente de las intenciones del conductor, mientras que otros únicamente actúan cuando se expresa la voluntad de efectuar un cambio de carril mediante el uso del intermitente.

Con el objetivo de minimizar las falsas alarmas, generalmente actúan por encima de un umbral de velocidad determinado y son capaces de realizar un filtrado de vehículos estacionados o de aquellos que circulan en sentido contrario. La zona de detección es unos 10 metros por detrás del espejo retrovisor por unos 4 de anchura, suficiente para cubrir el ángulo muerto.



Ilustración 11. Imagen vista por un espejo retrovisor

El primer vehículo en integrar el sistema de detección de objetos en ángulo muerto fue el Volvo XC90 (otoño 2005), utilizando para ello un par de cámaras CMOS integradas en los espejos retrovisores exteriores.

2.1.2 Asistente para cambios de carril

El asistente para cambios de carril es un sistema que amplía las posibilidades de funcionamiento del sistema de detección de objetos en ángulo muerto.

La distancia de detección aumenta hasta 50 ó 60 metros por detrás del vehículo y en los carriles adyacentes al mismo. Tiene en cuenta, además, la velocidad relativa del vehículo detectado en dicha zona con respecto al propio. De esta forma, se está en disposición de alertar al conductor en caso de existir un cierto riesgo al efectuar la maniobra de cambio de carril debido a la aproximación de otro vehículo a gran velocidad. En función de diversos parámetros, se pueden establecer diversos niveles de alerta.



Ilustración 12. Alerta en una aproximación lenta



Ilustración 13. Alerta en una aproximación rápida

El Audi Q7 fue el pionero en la utilización de un sistema de asistencia al cambio de carril, mediante el uso de dos radares de medio alcance ubicados en las esquinas del paragolpes posterior.

2.1.3 Sistema de visión nocturna

[7] Los sistemas de visión nocturna, basados en tecnología infrarroja permiten, mediante el uso de una pantalla integrada en el tablero del vehículo mejorar la visibilidad en condiciones de baja iluminación, permitiendo al conductor detectar transeúntes y otros riesgos potenciales que se encuentren en el camino.

De acuerdo con cifras presentadas por la compañía alemana Siemens, a pesar que durante la noche existe en promedio aproximadamente 80% menos tráfico que durante el día, alrededor del 33% los accidentes automovilísticos que cobran víctimas mortales, ocurren durante la noche.

El sistema de visión nocturna que permita detectar elementos difíciles o imposibles de detectar por el ojo humano a distancias de hasta 150m, dando al chofer el tiempo necesario para reaccionar ante esta situación.



Ilustración 14. Visión nocturna

La zona iluminada es grabada por una cámara instalada tras el parabrisas, y las imágenes son procesadas por una unidad de control y mostradas en la pantalla de alta

resolución del interior del coche. El sistema ahora es capaz de distinguir entre objetos fijos y móviles y cuando se detecta un peatón el conductor es convenientemente avisado en la pantalla, permitiendo reaccionar de forma adecuada y rápida.

2.1.4 Alerta por cambio involuntario de carril

[8] Es un sistema que alerta al conductor si abandona el carril por el que circula sin conectar antes los intermitentes, lo que se toma por una distracción. Se compone de un sistema capaz de detectar las líneas del carril por el cual se está circulando, una centralita electrónica y un sistema de aviso al conductor.

La detección de las líneas de carril puede realizarse mediante una serie de sensores de infrarrojos instalados en la parte inferior del vehículo, que utilizan la luz reflejada por las líneas de la calzada para detectar si el vehículo circula sobre éstas. En ese caso, si el conductor no ha activado los intermitentes, una centralita electrónica interpreta que se está abandonando involuntariamente el carril y alerta al conductor mediante diversos métodos: en unos casos hace vibrar el asiento, en otros el volante y en otros emite avisos sonoros y luminosos.

Otra forma de detectar las líneas que delimitan un carril consiste en el análisis de las imágenes provenientes de una cámara (instalada generalmente en el pie del espejo retrovisor interior). Este sistema presenta como ventaja fundamental la posibilidad de reaccionar ante una trayectoria conocida, pudiendo predecir la salida de carril antes de que ésta se produzca. Además, es un sistema que no genera falsas alarmas ante otras líneas, como son flechas pintadas en la calzada. Su funcionamiento en casos de visibilidad reducida es peor.



Ilustración 15. Sensores para detectar las líneas de carril

En todo caso, el sistema funciona únicamente a partir de una cierta velocidad (60 – 80 km/h es lo más habitual) y es desconectable. Asimismo, al activar el intermitente correspondiente, se interpreta que el conductor realmente desea realizar la maniobra de abandonar el carril por el que se circula y, por lo tanto, éste no es alertado.

Sus denominaciones comerciales son LDW (Lane Departure Warning) o AFIL (Alerte de Franchissement Involontaire de Ligne) en el caso de Citroën. Existe una evolución del sistema denominada LKAS (Lane Keeping Assistance System) en el cual se induce un par de giro a una dirección asistida eléctrica para indicar al conductor en qué sentido debe girar para mantenerse dentro de las marcas que delimitan el carril.

2.1.5 Control de crucero adaptativo (ACC)

[9] La tecnología de control de crucero adaptativo (ACC) mejora la función de control de crucero normal ajustando automáticamente la distancia y la velocidad en función de la marcha del vehículo que le precede. Esto se consigue con un detector de radar de gran alcance, un procesador de señales y un control longitudinal del vehículo. Si el vehículo que va delante reduce la velocidad o se detecta otro objeto, el ACC ajusta la velocidad del vehículo sin intervención del conductor. Cuando se haya despejado la carretera, el sistema volverá a acelerar el vehículo hasta la velocidad programada. Al igual que ocurre con el control de crucero normal, el conductor puede desactivar el

sistema cuando lo desee. Los actuales sistemas ACC son ante todo dispositivos de confort previstos para una gama de velocidades limitada.



Ilustración 16. Control de crucero adaptativo

Un ACC que funcione a cualquier velocidad resulta más seguro e incluso de manejo más cómodo. Si es preciso, el sistema desacelerará el vehículo hasta detenerlo completamente, empleando para ello toda su potencia de frenado, en lugar de desactivarse a una determinada velocidad, como ocurre en los sistemas actuales. El sistema también detecta si el coche de delante empieza a moverse de nuevo y avisa al conductor, generalmente por medio de una señal acústica. Entonces el vehículo acelera automáticamente hasta la velocidad programada, mantiene la distancia adecuada con el vehículo de delante y se adapta al ritmo del tráfico.

Algunas de las funciones de los sistemas ACC más avanzados son las siguientes:

- "Stop&Go".

Los sistemas ACC avanzados permiten que el vehículo, de forma automática y mediante un sensor, disminuya la velocidad al detectar un coche que circula más despacio por delante y recupere la marcha inicial programada cuando éste ya no se encuentre en la trayectoria.

- Reconocimiento de límites de velocidad.

El sistema de detección de señales de tráfico mediante cámaras es otra de las funciones del ACC avanzado. El conductor recibe la información sobre la velocidad

máxima permitida por medio de voz y visual a través del cuadro de instrumentos. Como opción, el coche también adapta su velocidad a la máxima que indica la señal de tráfico.

- Control para mantener la trayectoria.

Gracias a un sistema de reconocimiento de imágenes, es posible detectar de forma fiable los bordes del carril. En esta función, una unidad electrónica calcula la distancia entre el coche y las líneas del carril. En el futuro, un aviso acústico o una vibración del volante alertará al conductor en caso de abandonar la trayectoria.

- Control de velocidad en curva.

Una de las principales limitaciones de los actuales sistemas ACC se presenta al trazar curvas cerradas. En estos casos, el coche que marcha por delante puede quedar fuera del campo de visión del sensor y éste también puede reconocer un vehículo que circula en sentido contrario y que no es relevante en nuestra trayectoria. Estos inconvenientes se solventan mediante los dispositivos de reconocimiento de las líneas del carril y el sistema de navegación predictiva. Gracias a los primeros, el coche reconoce continuamente los bordes de la carretera, mientras que el segundo permite ajustar la velocidad del vehículo a las condiciones de la carretera mediante una serie de parámetros adicionales incluidos en el CD de navegación.

2.1.6 Sistemas de Control de la Presión de los Neumáticos

[10] Existen dos tipos principales de sistemas de control de la presión de los neumáticos: los sistemas de detección indirecta y los sistemas de detección directa. Al margen de éstos, existen otros tipos de sistemas alternativos que también pueden resultar prácticos a la hora de mejorar la seguridad en la conducción.

a) Sistema de detección indirecta de la pérdida de presión:

Este sistema asocia las pérdidas de presión a las reducciones de la circunferencia del neumático medidas a través de los sensores de velocidad de las ruedas. Su funcionamiento se basa en que los descensos de presión provocan una reducción del radio efectivo de la rueda, y por tanto una mayor velocidad de giro con respecto al resto de neumáticos (las ruedas con menor radio han de girar más rápido para recorrer la misma distancia). El sistema compara la velocidad relativa de una rueda con respecto a las demás y avisar al conductor cuando detecta diferencias significativas.

La principal ventaja de este sistema es que resulta muy fácil de implementar en un vehículo que esté equipado con sistema de antibloqueo de frenos (ABS), ya que la velocidad de rotación de las ruedas se puede obtener directamente de los sensores de velocidad de giro de las ruedas del propio sistema ABS. Tras un acuerdo alcanzado entre la Comisión Europea y los fabricantes de vehículos, todos los automóviles nuevos disponen en la actualidad de sistema ABS. Las desventajas radican en que este sistema no es normalmente capaz de decir qué rueda es la que está desinflada, ni de detectar una pérdida simultánea de presión en todas las ruedas, así como tampoco de comprobar la rueda de repuesto. Este sistema tampoco es capaz de detectar pérdidas de presión por debajo de un cierto umbral mínimo.

b) Sistema de medición directa de la presión en los neumáticos:

Con este sistema, mediante un pequeño sensor (o manómetro) en cada rueda, resulta posible detectar incluso pequeñas caídas de presión respecto a la recomendada por el fabricante del vehículo. La mayoría de estos sistemas disponen de sensores de temperatura y de presión acoplados a la válvula de inflado. Los datos son transmitidos mediante antenas hasta un módulo de control, el cual los analiza y muestra al conductor a través de un testigo o pequeña pantalla digital en el salpicadero.

El coste de un sistema de medida directa de la presión es superior al coste del sistema de detección indirecta, y además exige un cierto mantenimiento (por ejemplo, la sustitución de los sensores o las válvulas averiados).

Por otro lado, las ventajas con respecto al sistema indirecto son la mayor sensibilidad a los cambios de presión (detecta desviaciones de hasta 0,1 bares o kgf/cm² de presión) y la posibilidad de medir presiones individualmente en cada rueda, incluso en la de repuesto, y con el coche parado.

La siguiente ilustración muestra de modo esquemático el funcionamiento de un sistema de medida directa de la presión de los neumáticos en el mismo tapón de la válvula) mostrará un color verde, amarillo si la presión descende y de color rojo cuando su valor sea muy bajo.

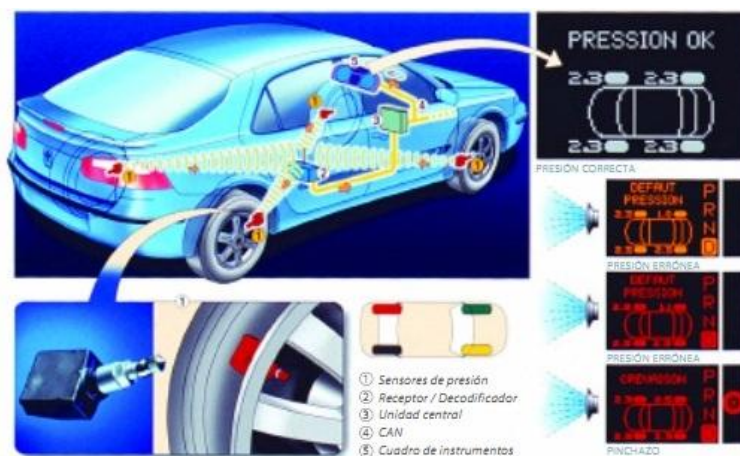


Ilustración 17. Sistema de detección de presión de neumáticos

2.1.7 Sistema de detección de señales

[11] Los sistemas de detección de señales constan de una cámara, situada entre el parabrisas y el espejo retrovisor interno, y dos procesadores de las imágenes que trabajan con las 30 imágenes por segundo que es capaz de detectar el sistema. Uno de

los procesadores está destinado al reconocimiento de las señales de circulación y otro para detectar las líneas que definen el carril y la calzada.

Traffic Sign Recognition es el sistema de Opel que controla las señales de circulación. Simplemente analiza las imágenes y reconoce las señales hasta una longitud de 100 metros y las muestra en el display del conductor. Detecta diferentes tipos de señales pero dota de mayor prioridad a señales tan importantes como el fin de la zona de adelantamiento.



Ilustración 18. Sistema de detección de señales

Capítulo 3. Estudio sobre la detección de peatones

En el siguiente capítulo vamos a realizar un estudio detallado sobre los distintos sistemas que existen en la actualidad para la detección y seguimiento de peatones en un entorno de conducción mediante técnicas de fusión sensorial, cada una de ellas distinta pero que tienen un mismo objetivo común, detectar peatones para así evitar una situación de riesgo y una posible colisión.

3.1 Sensor Fusion-Based Pedestrian Collision Warning System with Crosswalk Detection [12]

Autores: Shigetaka Suzuki, Pongsathorn Raksincharoensak, Ikuko Shimizu, Masao Nagai y Rolf Adomat. *Department of Mechanical Systems Engineering Tokyo University of Agriculture and Technology and Continental, Division Chassis and Safety.*

3.1.1 Introducción

En este método se va a utilizar una fusión sensorial de una cámara estéreo y un radar de onda milimétrica para obtener una alta fiabilidad del sistema. Para el desarrollo de sistema de protección de peatones frente a accidentes es importante medir con precisión la posición de los peatones. Este enfoque tiene la ventaja de una alta robustez, una baja tasa de falsas alarmas y una implementación rentable. El vehículo experimental detecta el peatón cerca del cruce de peatones, estima la probabilidad de una colisión, y por último avisa al conductor con advertencias en el caso de que se dé una situación de peligro.

En la ilustración 19 podemos observar un diagrama del sistema:

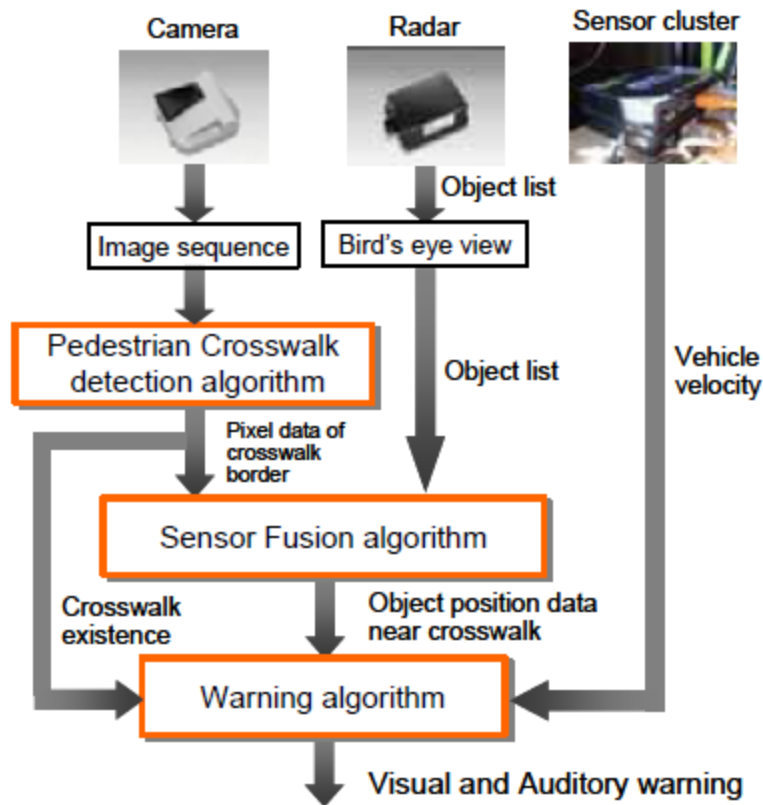


Ilustración 19. Diagrama de bloques del sistema de fusión.

3.1.2 Características del vehículo de pruebas

El vehículo de pruebas que se va a utilizar se equipará con un radar de onda milimétrica, una cámara monocular CMOS, un PC, y un grupo de sensores mediante los cuales se obtiene la información del vehículo así como la velocidad y la velocidad de derrape. El algoritmo que se va a utilizar permite segmentar los objetos detectados por el radar de ondas milimétricas y devuelve a la salida los datos de posición de los mismos gracias al seguimiento mediante filtros de partículas. Por lo tanto, la posición y la velocidad relativa de los objetos detectados se pueden obtener a través del radar de ondas milimétricas. Estos datos son datos de entrada a un PC a través de un dispositivo de bus CAN y puede ser monitorizado en tiempo real. La imagen de la cámara monocular es capturada por un interfaz de medición en tiempo real y la imagen se envía al PC para procesar las imágenes.

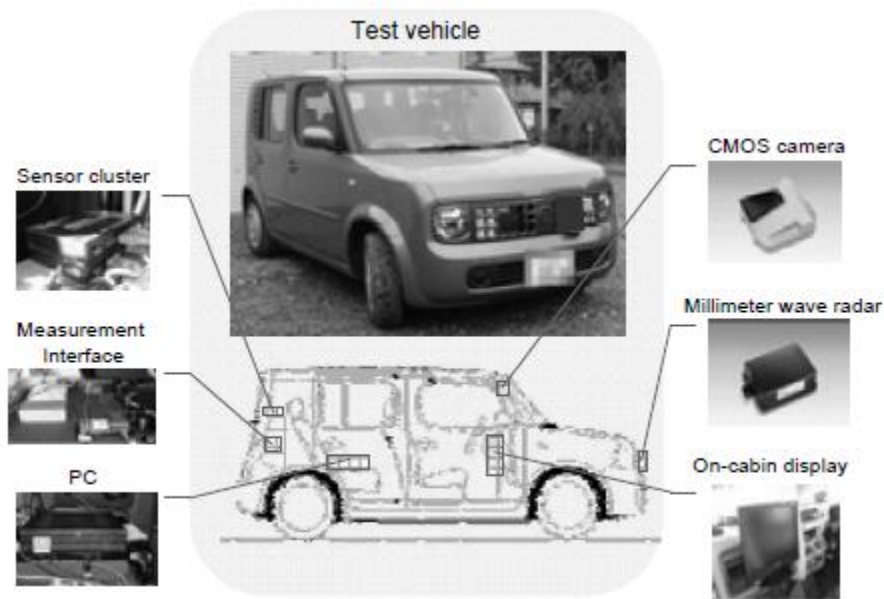


Ilustración 20. Vehículo de pruebas utilizado y colocación en el mismo de los sensores.

3.1.3 Diseño del sistema de detección de peatones y alerta

Una vez obtenida la imagen se procesa mediante tres algoritmos.

- *Algoritmo de detección en cruz*

El primero de ellos es un algoritmo de detección del paso de peatones. Dicho algoritmo utiliza una característica llamada “*cross ratio*” (relación en cruz) de cuatro puntos alineados. Ya que la relación cruzada es invariante bajo la proyección en perspectiva, la relación cruzada del paso de peatones en el espacio 3D es la misma que la de su imagen, como se muestra en la ilustración 21.

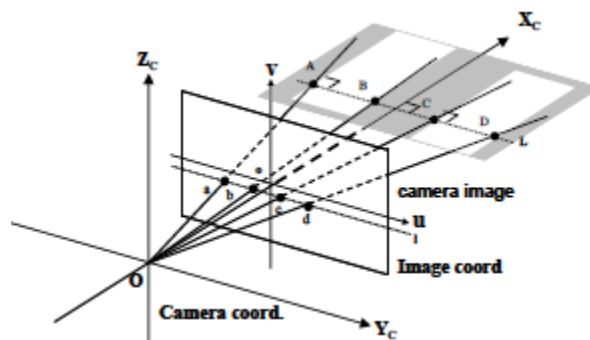


Ilustración 21. Descripción gráfica de la relación en cruz.

Aquí, la relación en cruz [abcd] de cuatro puntos a, b, c, d se calculan de la siguiente manera: en el símbolo escrito como ab se refieren a la distancia entre el punto a y el punto b.

$$[abcd] = \frac{\overline{ac}}{\overline{bc}} \cdot \frac{\overline{bd}}{\overline{ad}}$$

Ecuación 1

El diagrama de flujo del algoritmo de detección de peatones se muestra en la ilustración 22.

En primer lugar, para extraer los límites de las líneas blancas de una carretera, se procesa una imagen captada por la cámara mediante binarización y detección de bordes. Para extraer los bordes de la imagen de entrada, se utiliza el algoritmo de Canny. Entonces, los bordes verticales se clasifican en dos tipos: los límites en los cambios de intensidad de oscuridad a blanco o de blanco a oscuro. Por último, se calcula la relación cruzada de los dos pares de límites. Se extraen cuatro puntos de la imagen, que tienen la misma altura. En este paso, los candidatos del paso de peatones en la imagen se reducen mediante la comprobación del número de límites a la misma altura de la imagen y el ancho entre los dos bordes. A continuación, se calcula la relación cruzada [abcd] de cuatro puntos a, b, c, d, como se indica en la ecuación 1.

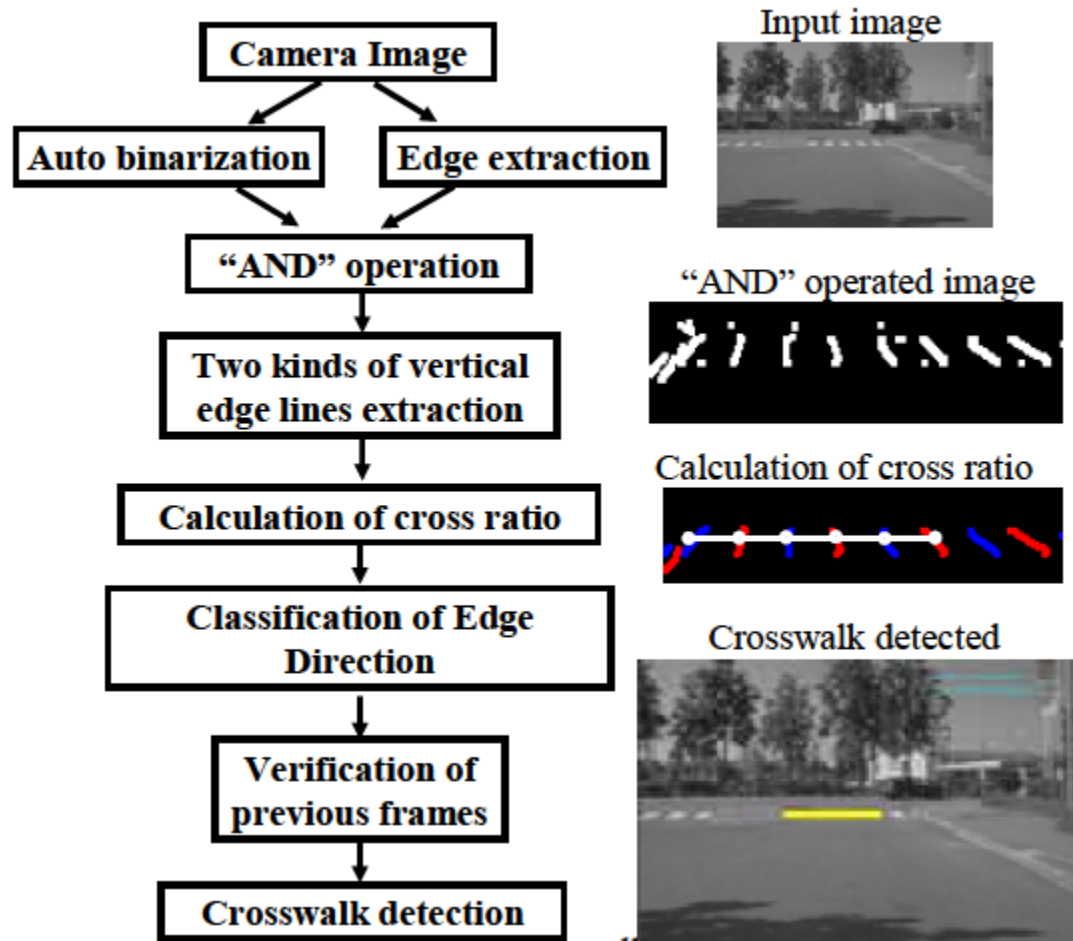


Ilustración 22. Diagrama de flujo del algoritmo de detección de peatones con su correspondiente imagen para los bloques principales.

Si los bordes de tres pares extraídos de los resultados del cálculo están ajustados a la cruz de la relación de paso de peatones, los bordes extraídos son considerados como un paso de peatones.

- *Algoritmo de fusión de sensores*

El siguiente algoritmo que se utilizará será un algoritmo de fusión de la información de varios sensores, cuyo diagrama de flujo se muestra a continuación en la ilustración 23:

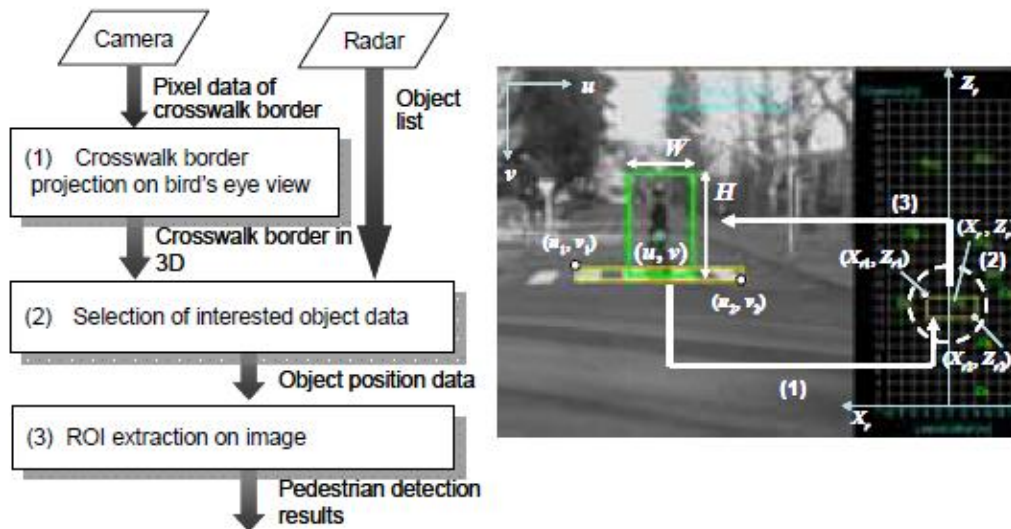


Ilustración 23. Diagrama de flujo del algoritmo de fusión de la cámara y el radar.

Lo primero será obtener las coordenadas en píxeles del límite del paso de peatones detectados a partir de la unidad de procesamiento de la imagen y los datos del objeto hacia la posición con respecto al vehículo de prueba obtenida del radar de onda milimétrica. Pasos:

- 1) Las coordenadas en píxeles de los límites del paso de peatones de la imagen de la cámara se proyectan en el sistema de coordenadas. La relación del sistema de coordenadas de la cámara (u, v) y el sistema de coordenadas del radar (X_r, Y_r, Z_r) se puede expresar de la siguiente manera en la ecuación 2:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X_r \\ Y_r \\ Z_r \\ 1 \end{pmatrix}$$

Ecuación 2

Donde p_{11} p_{12} p_{13} ... p_{34} son los elementos de la perspectiva de la matriz de proyección para la transformación de coordenadas. Los coeficientes de la matriz se determinan en función de la relación entre la posición de un

objeto de referencia, la información de las imágenes de la cámara y la posición medida con el radar (calibración). En el proceso de cálculo, Z_r se establece como una constante que corresponde a la altura del sensor con respecto a la altura del plano de la carretera.

2) Los objetos que se encuentran sobre o cerca de un cruce de peatones se seleccionan como candidatos de un peatón. En este caso, la región de paso de peatones se ha ampliado con un margen de 1 metro a cada lado de a fin de detectar a los peatones cerca del cruce de peatones.

3) Desde el paso anterior, los datos de radar (X, Z), que indican los candidatos de los peatones, se proyectan en el sistema de coordenadas de la cámara (u, v), y el grupo de esquinas de píxeles (u_i, v_i) están conectados como un rectángulo determinado por anchura W y la altura H . La caja rectangular se convierte en la región de interés (ROI).

- *Algoritmo de alerta*

Por último está el algoritmo de alerta que proporciona una representación visual y un sistema de alerta auditiva en tiempo real al conductor en función del riesgo de colisión. El algoritmo de advertencia estima que el riesgo de colisión de la posición predicha del peatón mediante un sistema de tiempo de colisión (TTC).

El diagrama se muestra en la siguiente ilustración:

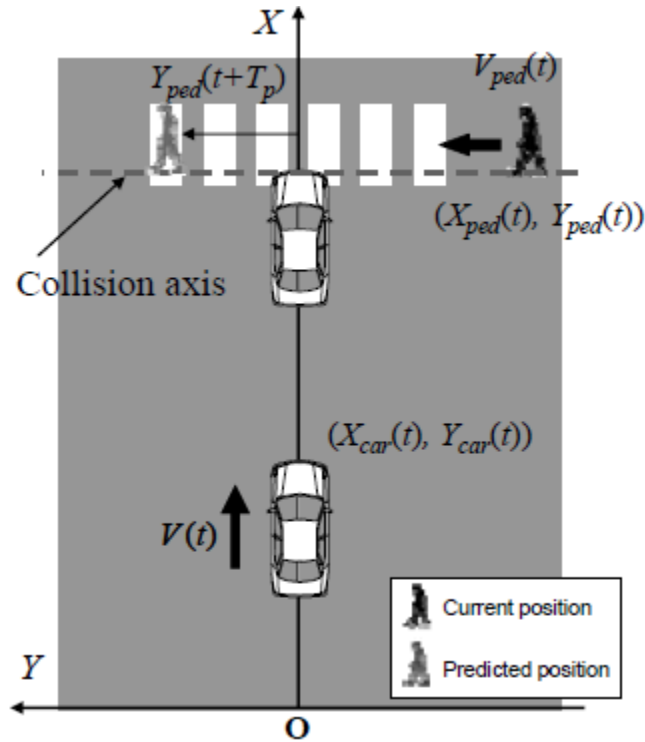


Ilustración 24. Descripción gráfica de la protección de los peatones previstos.

En primer lugar, se define la dirección del movimiento del peatón, y después se calcula el tiempo previsto (T_p) del vehículo para llegar a dicha posición. La posición predicha para los peatones es la posición en la que un peatón se mueve por el tiempo (T_p). Los siguientes son supuestos para el cálculo: Los peatones y los vehículos continuarán moviéndose con velocidad constante ($V_{ped} = \text{const}$, $V_{car} = \text{const}$). El vehículo de prueba se mueve en línea recta y perpendicular al paso de peatones ($Y_{car} = 0$). Los peatones se mueven a lo largo del eje móvil definido anteriormente con una velocidad media constante ($X_{ped} = \text{const}$). El tiempo previsto (T_p) de acuerdo con estas condiciones se expresa con la ecuación 3:

$$T_p(t) = \frac{X_{ped}(t) - X_{car}(t)}{V_{car}}$$

Ecuación 3

Y luego la posición de los peatones se puede predecir de la siguiente manera en la ecuación 4:

$$Y_{ped}(t + T_p(t)) = Y_{ped}(t) + V_{ped}(t) \cdot T_p(t)$$

Ecuación 4

Donde,

$X_{ped}(t) - X_{car}(t), Y_{ped}(t)$ son posiciones medidas por el radar. Por otro lado la velocidad de paso de peatones (V_{ped}) se calcula por el valor diferencial de la posición lateral (Y_{ped}).

La posición predicha de los peatones puede ser calculada en tiempo real desde la posición actual de peatones y la velocidad del vehículo adquiridos por los sensores. La región de interés para la detección se determina como se muestra en la ilustración 25:

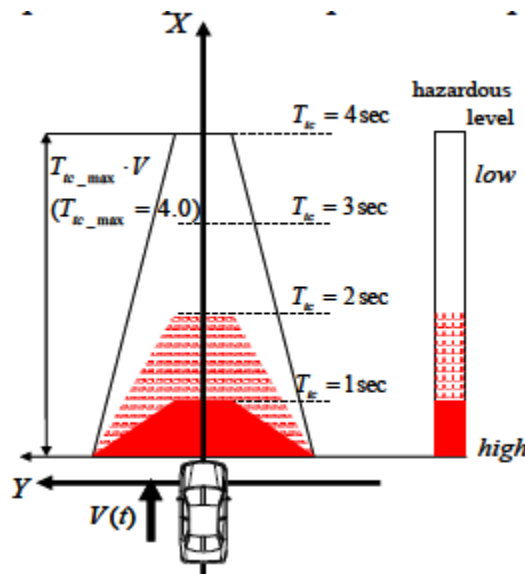


Ilustración 25. Definición de la región de interés y el nivel de riesgo

El borde superior de la región de interés se fija para ser proporcional a la velocidad del vehículo, T_{tc} indica el parámetro para establecer el rango longitudinal en la parte superior de detección de peatones que es equivalente al tiempo de colisión, y V_{car} indica la velocidad del vehículo. La relación proporcional entre el vehículo y el rango longitudinal T_{tc} se establece en el valor constante de 4,0

segundos. A partir de este valor se establece la peligrosidad de colisión de peatones por cada 1,0 segundos. Después, para comprobar los datos de salida de procesamiento del sistema, como se puede observar en la ilustración 18, se utiliza un semáforo en rojo / amarillo / verde dentro del vehículo:

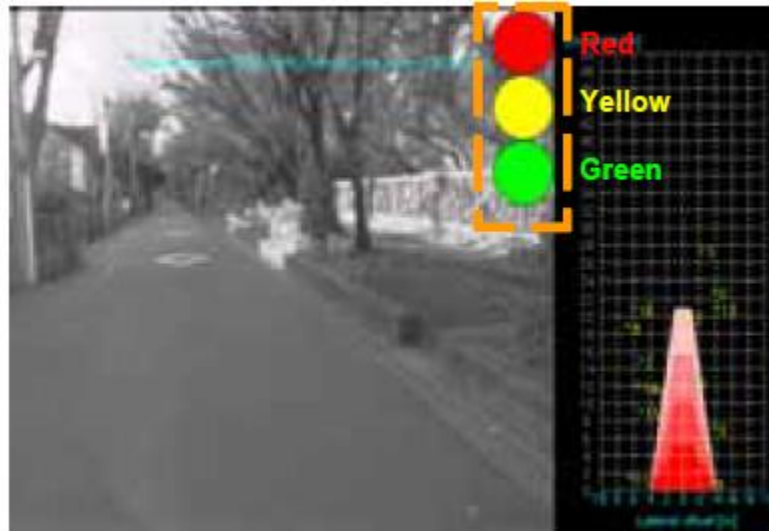


Ilustración 26. Información presentada en el *display*.

Dependiendo de la condición de la posición de los peatones y su velocidad al caminar, la advertencia de colisión consiste en una señal auditiva, y además el color de la advertencia visual en la pantalla del vehículo cambia de acuerdo a las siguientes reglas, y se expedirá un aviso auditivo al conductor cuando la situación se vuelva más crítica:

- a) Si no se detecta ni paso de peatones ni peatones, el color es verde.
- b) Si se detecta un paso de peatones o peatones cerca del vehículo el color es amarillo y se muestra en la pantalla "Cruce Detectado".
- c) Si se detecta un paso de peatones y peatones en movimiento en el paso de peatones y la posición predicha está por debajo del valor de $V_{car} \cdot T_{tc}$, el color es rojo y se muestra en la pantalla, "Cruce ocupado" y además si el valor de T_{tc} es inferior a 2,0 segundos, se proporciona al conductor un aviso auditivo, en las siguientes ilustraciones se muestra un ejemplo gráfico:

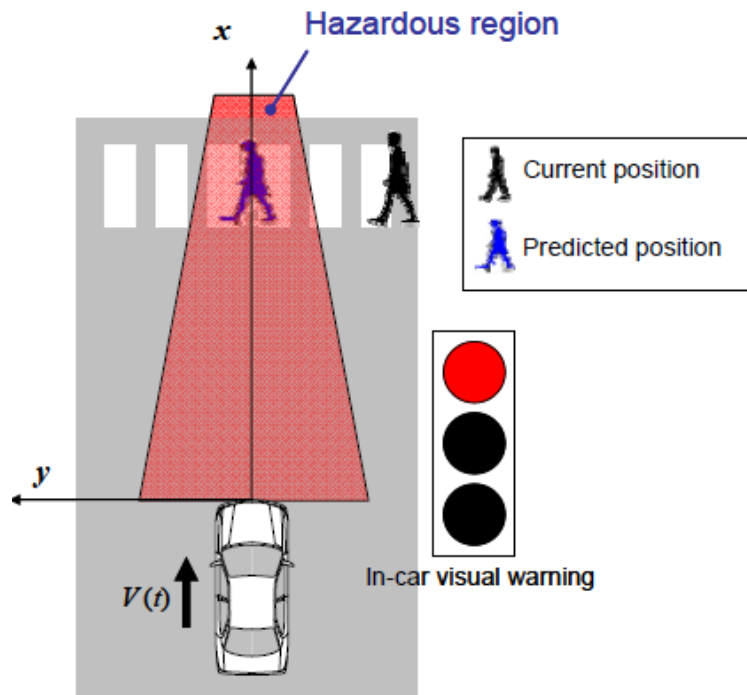


Ilustración 27. Advertencia de colisión de peatones cuando la situación es crítica.

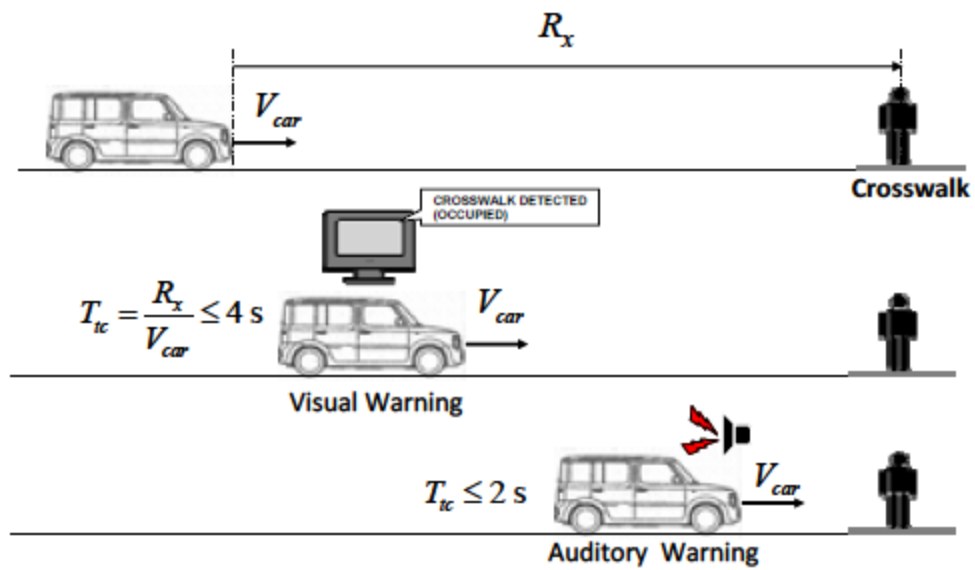


Ilustración 28. Descripción de las interfaces de advertencia de colisión de peatones.

3.1.4 Resultados experimentales

Una vez realizado el experimento con éxito se verifica la validez del método utilizado. Para ello los experimentos en el vehículo de prueba se llevaron a cabo para confirmar si el sistema para evitar colisiones con los peatones puede funcionar de manera satisfactoria. En las pruebas el vehículo se mueve con una velocidad constante de 20 km / h en perpendicular a un paso de peatones. Como se muestra en la ilustración 21, la condición experimental fue que un peatón empieza a caminar en el paso de peatones cuando el tiempo de colisión T_{tc} es menos de 3,0 segundos y con el vehículo en movimiento.

El resultado experimental se indica en la ilustración 22. Cuando el peatón llega a la región de interés, la luz artificial de tráfico en la pantalla se cambia de amarillo a rojo. En la ilustración 23 se muestra la trayectoria de los peatones con respecto a la posición del vehículo durante el experimento. De acuerdo con la ilustración 23, la trayectoria del peatón llega a la región de interés (región de advertencia). Por lo tanto, se puede verificar la validez de la propuesta de detección de peatones y el algoritmo del sistema de alerta.

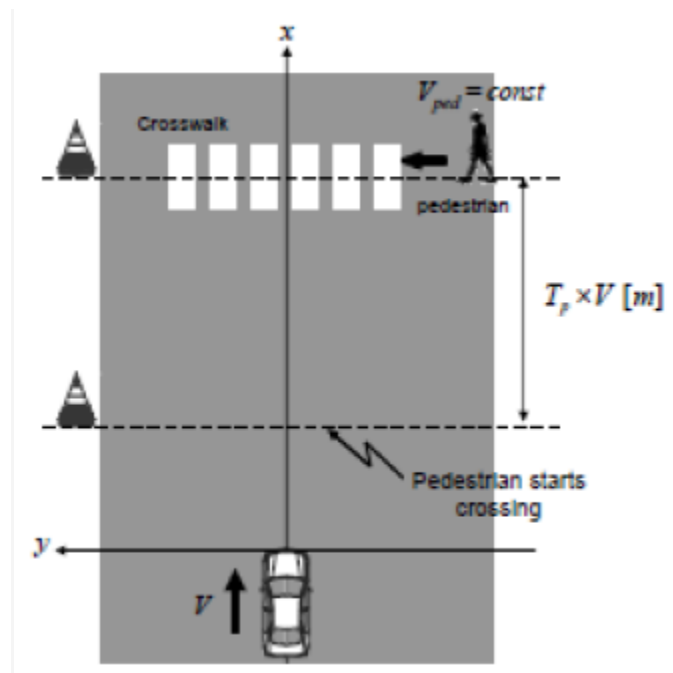


Ilustración 29. Condiciones experimentales.

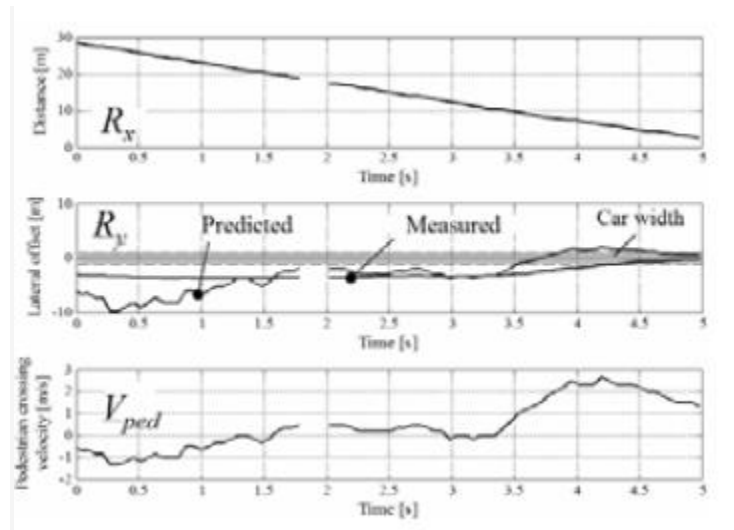


Ilustración 30. Resultado experimental de la detección de peatones

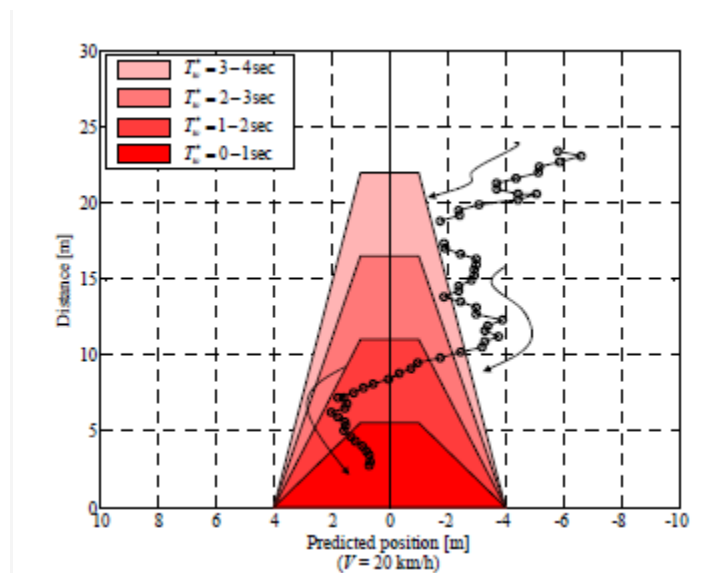


Ilustración 31. Tránsito de la posición predicha para peatones (a vista de pájaro)

3.2 Hybrid Fusion Scheme for Pedestrian Detection based on Láser Scanner and Far Infrared Camera. [13]

Autores: F. García, D. Olmeda, J. M. Armingol y A. de la Escalera. Intelligent System Lab, Universidad Carlos III de Madrid.

3.2.1 Introducción

Este método combina la información proporcionada por un telémetro láser 2D y una cámara de visión infrarroja lejana para detectar peatones en el entorno de la carretera. Aprovechando las diferentes capacidades de cada sensor y rangos de detección, cada detección de peatones se realiza de forma independiente. En las regiones donde los rangos se superponen, se utiliza un algoritmo de fusión para mejorar la detección y evitar los falsos positivos. Para ello se utilizará fusión tanto de bajo nivel como de alto nivel con los datos obtenidos por los sensores y así realizar una correcta detección y clasificación de los peatones.

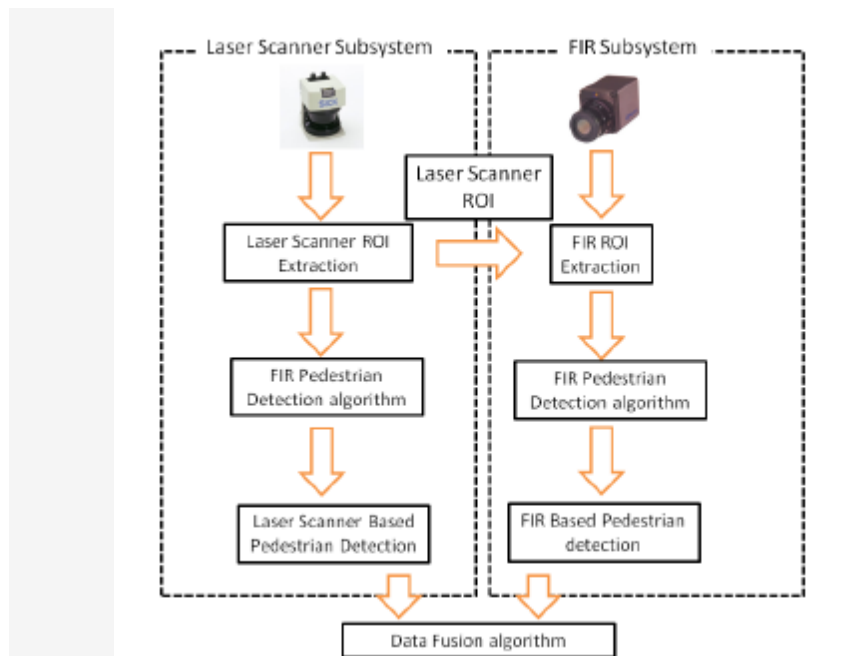


Ilustración 32. Diagrama de flujo de todo el sistema.

Como se muestra en la ilustración 32, se tienen dos subsistemas de detección, donde cada uno de ellos detecta los peatones por separado con diferentes sensores, tanto el escáner láser como el sistema de visión de infrarrojo lejano.

3.2.2 Algoritmo del subsistema del láser

El algoritmo del subsistema del láser consiste en la detección de bajo nivel y la identificación, no obstante también se calcula información de seguimiento. De acuerdo con el movimiento de los obstáculos y su posición, se pueden predecir las siguientes detecciones. El algoritmo se compone de dos etapas como se puede ver a continuación en la ilustración 33:

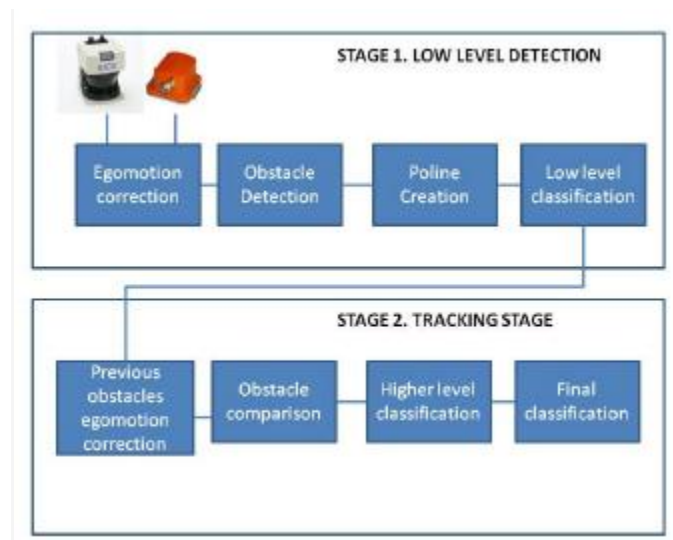


Ilustración 33. Procesos involucrados en el algoritmo de detección y clasificación usando el escáner láser.

3.2.3 Detección de bajo nivel

En la primera etapa, se reciben los datos y se realiza la identificación de bajo nivel. En la segunda etapa, los datos se integran en el tiempo y por lo tanto se obtiene un nivel de clasificación más alto. Esta etapa utiliza la información proporcionada por

las detecciones anteriores para predecir el movimiento de los peatones y así dar una clasificación más robusta.

La detección de bajo nivel consta de cuatro pasos:

- Corrección del movimiento del vehículo.

Los datos recibidos por el láser se corrigen de acuerdo al movimiento del vehículo. La información del movimiento del vehículo se registra por un sensor GPS incluido en el sistema. Este sensor GPS proporciona datos de posicionamiento e información como velocidad, aceleración, ángulos de Euler (velocidad absoluta y angular), etc.

- Detección de obstáculos.

Los puntos resultantes se unen de acuerdo a la distancia entre ellos después de la corrección del movimiento del vehículo. Se agrupan de acuerdo a un algoritmo basado en la distancia euclídea. Mediante el algoritmo se pueden corregir los segmentos ya creados, si llega el caso en que un punto no está incluido dentro de un segmento, se crea un nuevo segmento y los segmentos creados con un sólo punto se eliminan al considerarse falsos positivos.

- Creación de polilíneas.

Una vez que los segmentos se crean, los puntos contenidos dentro de cada segmento se combinan con líneas conocidas como polilíneas. Estas líneas se combinan entre sí de acuerdo con la distancia entre los puntos incluidos en ellas. El primer y el último punto se fusionan con una línea, por cada punto contenido dentro de este segmento, se calcula la distancia a la línea y si es superior a un umbral establecido se crean dos nuevas líneas fusionando estos tres puntos. Este proceso se repite para cada punto dentro de estas nuevas líneas.

- Clasificación de bajo nivel.

La clasificación de bajo nivel se realiza con la información proporcionada por las polilíneas de acuerdo con las formas de los diferentes obstáculos. Aquí pueden ser diferenciados los diferentes obstáculos: pequeños obstáculos, bordes de la carretera, posibles vehículos, obstáculos en forma de L, peatones y otros obstáculos. Lo que interesa son los peatones. El algoritmo utilizado para la detección de un posible peatón se basa en la detección de dos piernas. El movimiento de un peatón se modela con cinco situaciones posibles representadas a continuación en las siguientes ilustraciones:

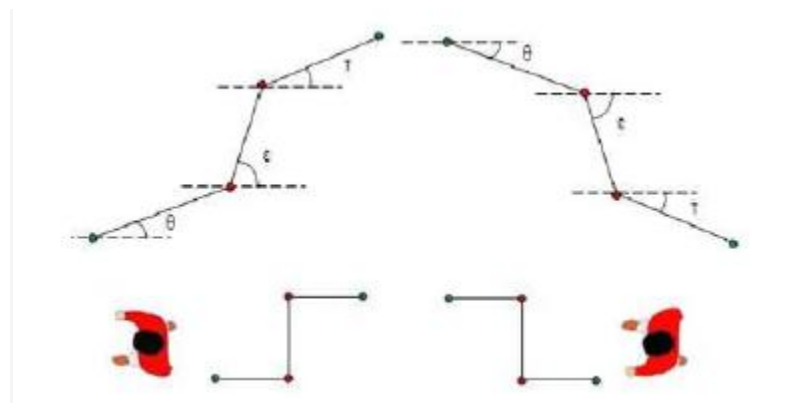


Ilustración 34. Modelos para movimientos perpendiculares al láser.

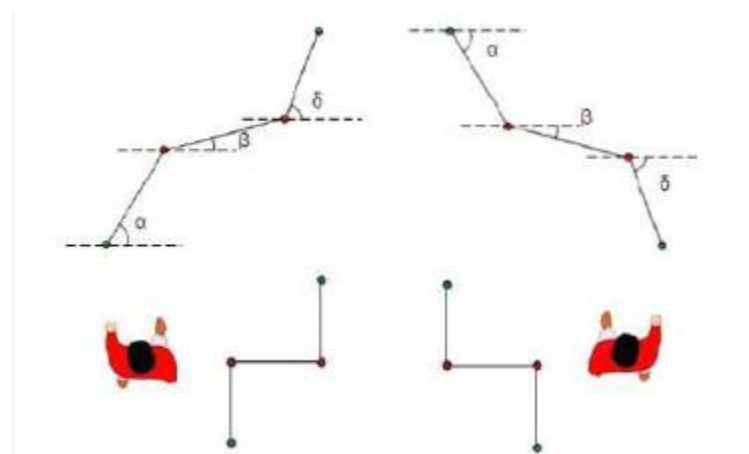


Ilustración 35. Modelos para movimientos paralelos al láser.

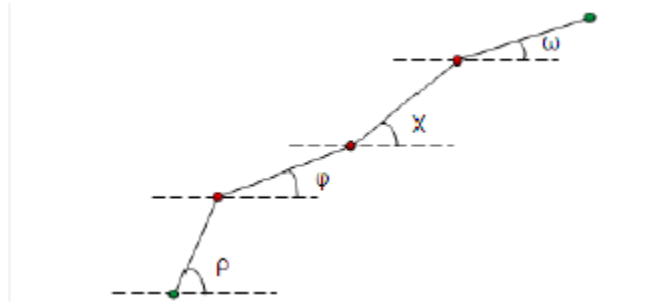


Ilustración 36. Modelo general para más de tres polilíneas.

3.2.4 Clasificación de alto nivel

A continuación se realiza una clasificación alto nivel, la cual consta de los siguientes pasos:

- Corrección del movimiento del vehículo respecto a un obstáculo previo.

En este paso se calcula el movimiento de los peatones detectados en los análisis anteriores y se prevé la siguiente posición para cada obstáculo. Los obstáculos se buscan dentro de una ventana de acuerdo con el tamaño de los obstáculos en análisis anteriores. Si un obstáculo se encuentra dentro de esta ventana de la exploración actual, se utiliza un algoritmo de comparación para verificar si el obstáculo es el mismo.

- Comparación de obstáculos.

El proceso de comparación se lleva a cabo de acuerdo con las características de forma, es decir, la anchura y la posición. Si hay varios posibles candidatos, el que tenga el valor más cercano es considerado como el mismo obstáculo.

- Clasificación de alto nivel.

Una vez que se ha llevado a cabo la detección de peatones de bajo nivel, en los siguientes ciclos, un obstáculo similar al ya detectado previamente que

pueda representar un pequeño obstáculo o un peatón, se considera un peatón si satisface previamente el algoritmo de comparación de acuerdo con algunas restricciones en relación con el movimiento y el tamaño. Para cada obstáculo, se almacena un registro de las últimas diez detecciones de bajo nivel para dar un mayor nivel de detección.

3.2.5 Subsistema de infrarrojo lejano

3.2.5.1 Modelo de cámara

En cuanto al subsistema de infrarrojo lejano, el sistema propuesto hace uso de imágenes en bruto de un microbolómetro no refrigerado, el nivel de gris de cada píxel de estas imágenes tomadas por la cámara representa la cantidad de calor que el sensor captura.

La homografía del plano de masa en el sensor se calcula para cada fotograma para determinar la posición. La proyección de un punto 3D en el plano de la imagen se puede hacer si se conoce su posición en relación a un plano determinado. La posición de la cámara en relación con el suelo se considera constante y se conoce gracias a una calibración previa. Los resultados de la proyección de un punto en el plano de la imagen son especialmente sensibles a las variaciones del ángulo de inclinación por lo que es necesario actualizar esta medida en cada fotograma. En el sistema que se presenta se conoce este punto de vista de cada cuadro de captura con la ayuda de un acelerómetro de tres grados de libertad sujeto a la base de la cámara.

3.2.5.2 Algoritmo de detección de peatones

Para comprobar la región de interés o ROI se utiliza como umbral la temperatura normal del cuerpo humano. El resultado es una imagen binaria, con manchas que pueden representar las partes del cuerpo humano con una diferencia alta de

temperatura, especialmente en la cabeza y las manos. En la primera fase se buscan las cabezas humanas, por lo tanto sólo se tienen en cuenta manchas en la mitad superior de la imagen. Una vez que las cabezas de los candidatos han sido seleccionadas, se genera una primera serie de regiones de interés y se crea un cuadro delimitador incluyendo todo el cuerpo del peatón. En la ilustración 37 se puede observar lo explicado anteriormente:

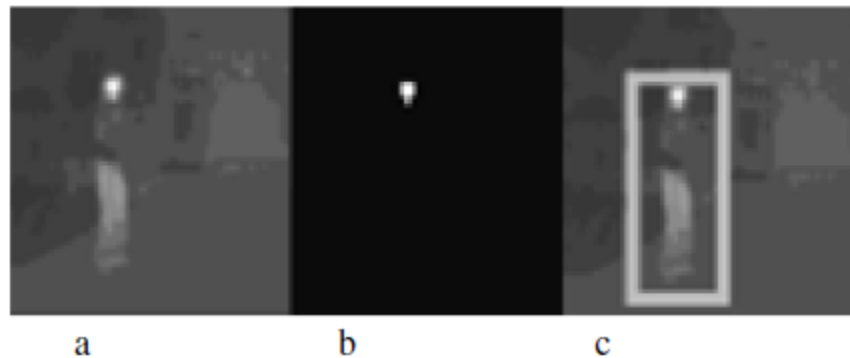


Ilustración 37. (a) Imagen recibida de la cámara. (b) Imagen binarizada con una cabeza. (c) Región de interés seleccionada.

Después se crea un segundo conjunto de la región de interés utilizando la información proporcionada por el láser. Si esta región no tiene correspondencia con la región del infrarrojo lejano detectado, se crea una nueva región. Con el láser se puede proporcionar la información de la anchura pero no de la altura.

Las regiones de interés generadas se digitalizan con un umbral que es la temperatura más baja establecida para el cuerpo humano. De esta forma se puede obtener la silueta de los posibles peatones incluidas en ella. La etapa de seguimiento se realiza mediante un filtro de Kalman.

Por último para realizar la fusión, se debe hacer una asociación de coordenadas. Se va a tomar como referencia el sistema de coordenadas del láser. La calibración de la fusión de sensores se realiza usando el algoritmo *Least Mean Square* (LMS). El algoritmo de calibración muestra una secuencia con un solo peatón. Este peatón es detectado por los dos subsistemas por separado y se utiliza un sistema de mínimos

cuadrados para obtener los coeficientes para convertir las coordenadas del sistema de coordenadas de la cámara de visión al sistema de coordenadas del láser.

3.2.6 Resultados

Los resultados obtenidos para cada subsistema tienen un porcentaje de detección del 90% para detección de peatones con el escáner láser en distancias inferiores a 30 metros. En las regiones de fusión los resultados muestran que más del 96% de los peatones han sido detectados.

3.3 A New Approach to Urban Pedestrian Detection for Automatic Braking [14]

Autores: *Alberto Broggi, (Miembro Senior de IEEE), Pietro Cerri, (Miembro de IEEE), Stefano Ghidoni, (Miembro de IEEE), Paolo Grisleri, (Miembro de IEEE) y Ho Gi Jung, (Miembro de IEEE). Vislab. Universidad de Parma.*

3.3.1 Introducción

En este apartado se describe un sistema para la detección de peatones sobre la base de un nuevo enfoque. Está diseñado para trabajar en un escenario urbano particularmente difícil. El sistema que se presenta se centra en un determinado escenario urbano en el que no sólo la detección de un peatón es de importancia fundamental, además el peligro de la situación también tiene que ser evaluado claramente.

En lugar de detectar todos los posibles candidatos y el filtrado de ellos después del análisis del entorno, en primer lugar se analiza el escenario y luego se realiza la búsqueda de posibles peatones en posiciones específicas para dichos escenarios. De esta manera, todos los peatones detectados corresponden a las posibles amenazas.

Este método se centra en la detección de peatones que aparecen justo por detrás de los obstáculos de oclusión, en la ilustración 30 se pueden observar ejemplos de peatones en zonas ocluidas:

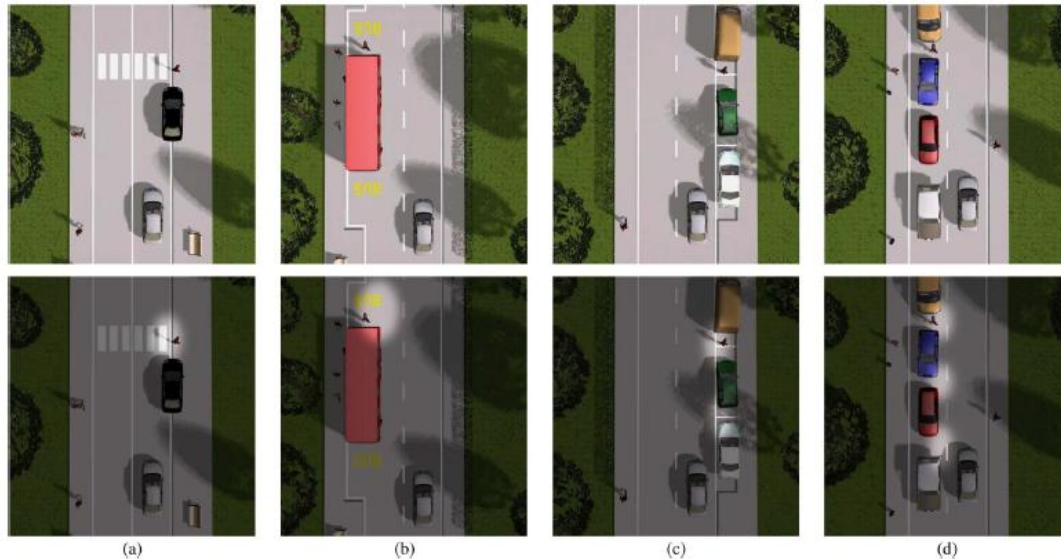


Ilustración 38. (Primera fila) Situaciones consideradas. (Segunda fila) Áreas de interés consideradas para que la detección de un peatón sea posible. (a) Un cruce peatonal está oculto por un vehículo estacionado. (b) Un peatón está cruzando la calle detrás de un autobús detenido. (c) Un peatón aparece entre dos vehículos estacionados. (d) Un peatón está cruzando la calle entre dos vehículos detenidos en el otro lado de la carretera. Es importante tener en cuenta que las situaciones (a) y (b) se refieren a determinados lugares establecidos en zonas urbanas (pasos de cebra y paradas de autobús). Situaciones (c) y (d) puede ocurrir en cualquier parte de la red de carreteras.

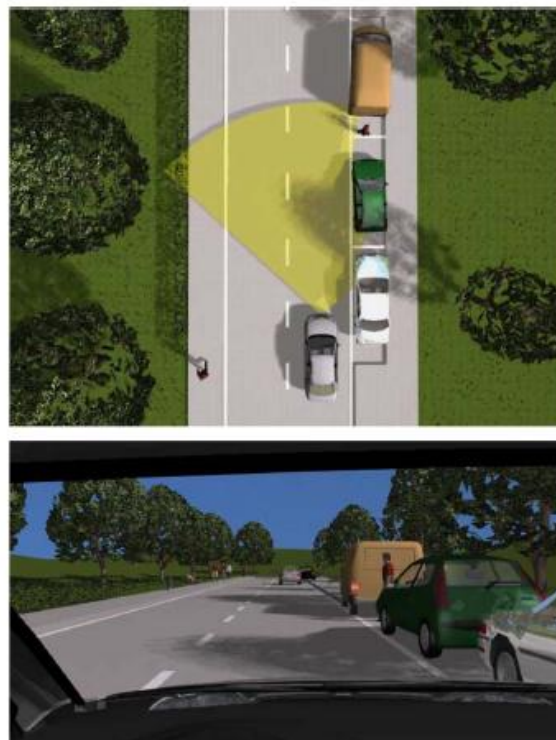


Ilustración 39. Peatón parcialmente oculto por los coches aparcados que no puede ser detectado por un escáner láser colocado en el parachoques delantero, pero puede ser detectado mediante la visión, aunque esté parcialmente ocluido.



Ilustración 40. Detalle de la defensa delantera que muestra la integración de escáner láser y los faros.



Ilustración 41. Detalle de cómo la cámara está instalada dentro de la cabina.

La principal característica requerida por este sistema es la capacidad para llevar a cabo lo siguiente:

- 1) Detectar rápidamente los peatones, dado el corto alcance y el peligro particularmente alto de colisión con un peatón que aparece de repente detrás de un obstáculo;
- 2) Detectar peatones tan pronto como aparecen, aun cuando siguen siendo parte ocluida;
- 3) Limitar la búsqueda a áreas específicas, que están determinadas por un preprocesamiento rápido.

La fusión del escáner láser y la cámara de visión puede proporcionar una rápida y sólida detección en caso de que de repente aparezcan peatones: El escáner láser proporciona una lista de áreas en las que un peatón puede aparecer, mientras que la cámara es capaz de detectar al peatón, incluso cuando todavía no es visible para el escáner láser.

3.3.2 Foco de atención.

Para hacer el algoritmo robusto, se han desarrollado dos clasificaciones diferentes de los datos del láser. Ambos métodos se basan en la agrupación impulsos en líneas rectas, pero mientras el primer método considera el último análisis sólo para clasificar los obstáculos, el segundo método utiliza una serie de análisis posteriores para verificar las posiciones de obstáculos y por lo tanto, estimar su velocidad.

Pasos:

A. Alineación de Datos

Los cambios posteriores entre mediciones con el escáner láser pueden ser apreciables y pueden causar problemas de agrupamiento o clasificación, ya que las formas de los obstáculos pueden aparecer distorsionadas.

Con los datos del movimiento del vehículo es posible estimar la rotación y traslación del vehículo y por lo tanto, corregir la posición medida por cada pulso. La velocidad del vehículo se utiliza para la traducción correcta, mientras que el ángulo del volante permite calcular la velocidad de giro o guiñada y, por tanto, la corrección de la matriz de rotación. La velocidad de giro puede ser computada como:

$$velocidad\ de\ giro = \frac{V_{fr} + V_{fl}}{t_w \cos \delta}$$

Ecuación 5

Donde V_{fr} y V_{fl} son las velocidades de la parte frontal derecha e izquierda respectivamente; δ indica el ángulo de las ruedas, y t_w es la vía delantera del vehículo. Como las medidas de guiñada se ven afectadas por el ruido, un sensor de guiñada adicional está instalado en el vehículo para aumentar la precisión y realizar una corrección de los datos más nítida.

B. Agrupación de datos

En primer lugar, los pulsos que pertenecen a la misma rotación están conectados juntos. Después, los obstáculos en movimiento son identificados por cuatro diferentes grupos y en paralelo, mientras que los obstáculos se interponen identificados por cuatro grupos superpuestos. Los pulsos se agrupan como cadenas de segmentos. Los puntos que no se pueden unir en cualquier cadena de la misma rotación, se revisan a la fusión con puntos de las rotaciones, considerando solamente la proximidad. Los puntos que no se puede conectar a cualquier otro punto o cerca del límite del rango del escáner láser se descartan.

C. Fusión del segmento

Hasta ahora, los pulsos se unen a las cadenas de segmentos sin cualquier otra información. Los segmentos adyacentes de aproximadamente la misma orientación se pueden combinar en un segmento más largo, preservando la forma de los obstáculos, pero reduciendo tanto la complejidad de la estructura de datos como los detalles de la representación. Teniendo en cuenta una cadena de segmentos, es posible calcular la línea recta que conecta el inicio de la cadena y el punto final. Entonces, se calcula la distancia entre esta línea y todos los puntos internos, y si la distancia máxima es mayor que el umbral, la línea es dividida en dos líneas. Estos pasos son iterados mientras que la distancia máxima es mayor que el umbral. El resultado es que cada cadena es por lo tanto, finalmente, segmentada en una polilínea.

La ilustración 42 muestra un vehículo en movimiento:

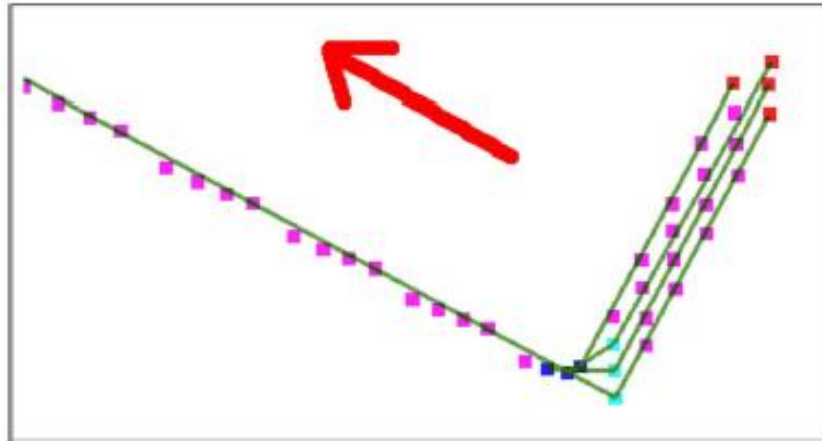


Ilustración 42. Un vehículo en movimiento y su dirección de movimiento. Las cuatro líneas paralelas corresponden a su parachoques trasero, mientras que la única línea corresponde a su lado izquierdo. Los puntos azules son los puntos de inicio de polilínea, los puntos de cian son los puntos de esquina, los puntos rojos son los puntos finales de polilínea, y los puntos de color violeta son puntos genéricos de línea.

El parachoques trasero se enmarca como cuatro líneas paralelas, mientras que su lado, está marcado por una sola línea que es paralela a la dirección del movimiento del vehículo.

D. Fusión de la línea.

Todos los obstáculos detectados por el escáner láser se identifican por cuatro líneas, es decir, una por cada rotación de láser. En el caso de obstáculos estáticos, las cuatro líneas son casi perfectamente superpuestas, debido a la corrección de datos anterior, y se pueden combinar. Por el contrario, en el caso de obstáculos en movimiento, las rectas son paralelas, pero no se superponen; por lo tanto, los obstáculos estáticos y en movimiento pueden ser localizados y clasificados correctamente.

E. Primera clasificación de obstáculos

Las polilíneas obtenidas hasta el momento se pueden clasificar de acuerdo sólo a su tamaño y forma. Es posible dividir los obstáculos en las siguientes categorías:

- 1) posibles peatones;
- 2) infraestructura vial;
- 3) obstáculos en forma de L;
- 4) obstáculo genérico.

Los obstáculos que se clasifican como peatones se supone que tienen un tamaño limitado mientras que los elementos viales son de un tamaño grande. Los obstáculos en forma de L se detectan mediante un método simple y eficaz. Todos los obstáculos que aún no se han clasificado son etiquetados como obstáculos genéricos. En la ilustración 35 se muestra un ejemplo de ello:

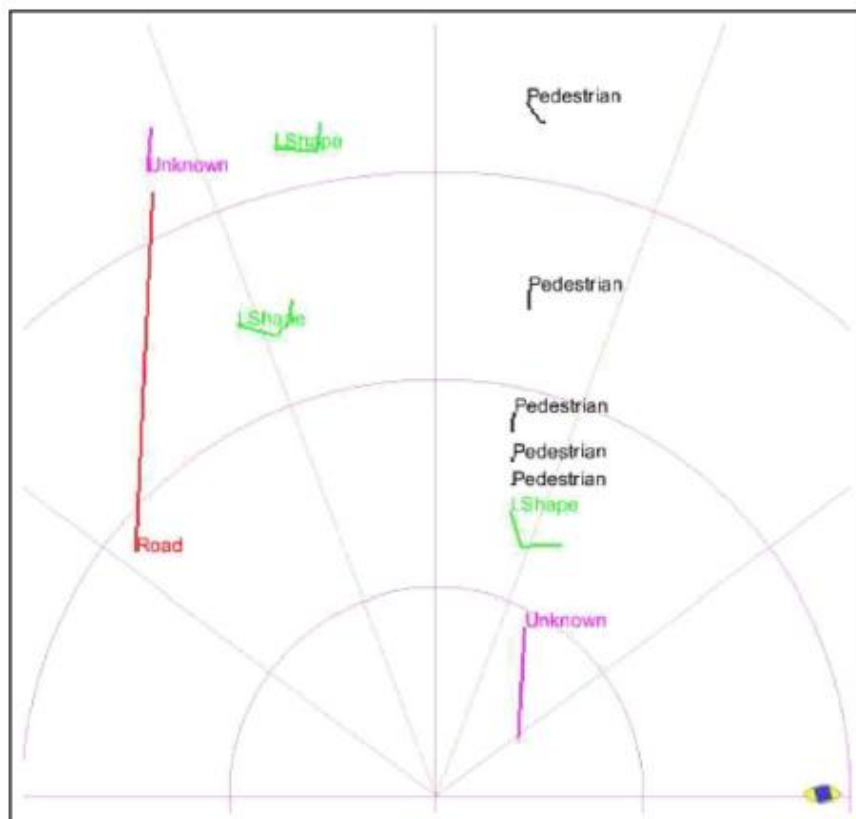


Ilustración 43. Clasificación de obstáculos. Todas las clasificaciones posibles basadas en la forma están representadas en la imagen.

F. Correlación temporal y basada en la forma.

La clasificación explicada en la sección anterior clasifica todos los pequeños obstáculos como posibles peatones. Incluso aunque es posible que todos los peatones estén correctamente clasificados, pueden estar presentes una serie de falsos positivos tales como obstáculos fijos a lo largo de la carretera (vehículos estacionados en particular). Por ello se establece una nueva clasificación de acuerdo a las siguientes cuatro clases:

- 1) obstáculos en movimiento,
- 2) obstáculos estáticos,
- 3) cambio en la forma del obstáculo,
- 4) nuevo obstáculo.

G. Zona de conducción

Los obstáculos identificados como estáticos en la segunda fase de clasificación se utilizan para definir la estructura del entorno. Los obstáculos fijos se utilizan para construir la zona de conducción, es decir, el área que puede ser alcanzada por el vehículo en un futuro próximo, como se muestra en la ilustración 36.

La zona de conducción puede estar compuesta de varias bifurcaciones, pero el ancho de cada subzona debe ser mayor que el ancho del vehículo. Es importante señalar que esta zona no siempre coincide con el espacio libre en la parte delantera del vehículo, ya que se construye sólo con obstáculos fijos y también puede incluir el desplazamiento de obstáculos.

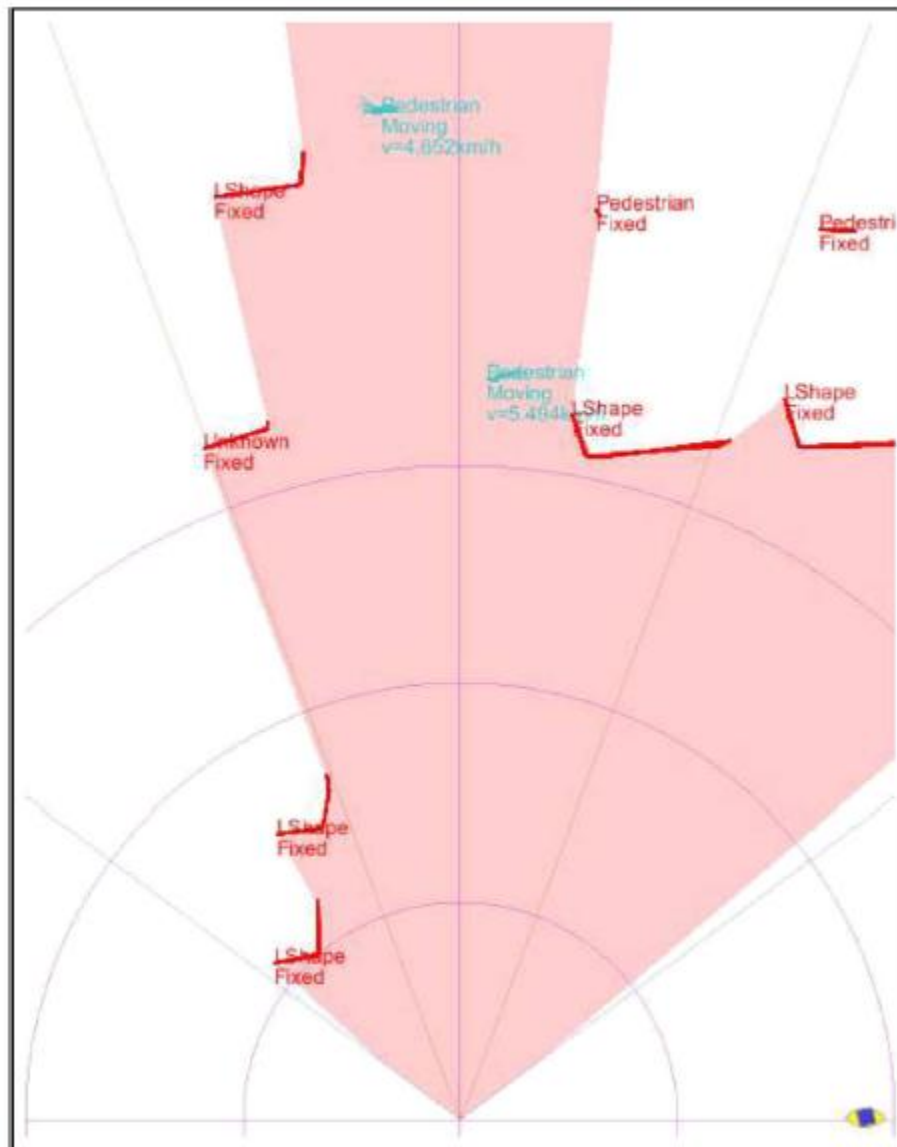


Ilustración 44. Clasificación de obstáculos. Se muestran todas las posibles clasificaciones basadas en el movimiento. Las polilíneas en movimiento cambian su posición en el sistema de coordenadas. La zona de conducción está destacada en color rosa.

H. Áreas peligrosas y posibles peatones.

Estas áreas están ubicadas a lo largo de los bordes de la zona. Los obstáculos estáticos que se utilizan para construir la zona de conducción también se utilizan para identificar las áreas peligrosas, es decir, áreas en las que los peatones pueden aparecer. Las áreas peligrosas se encuentran detrás del punto más alejado de cada polilínea estática, como se muestran en la ilustración 37 representadas por círculos rojos:

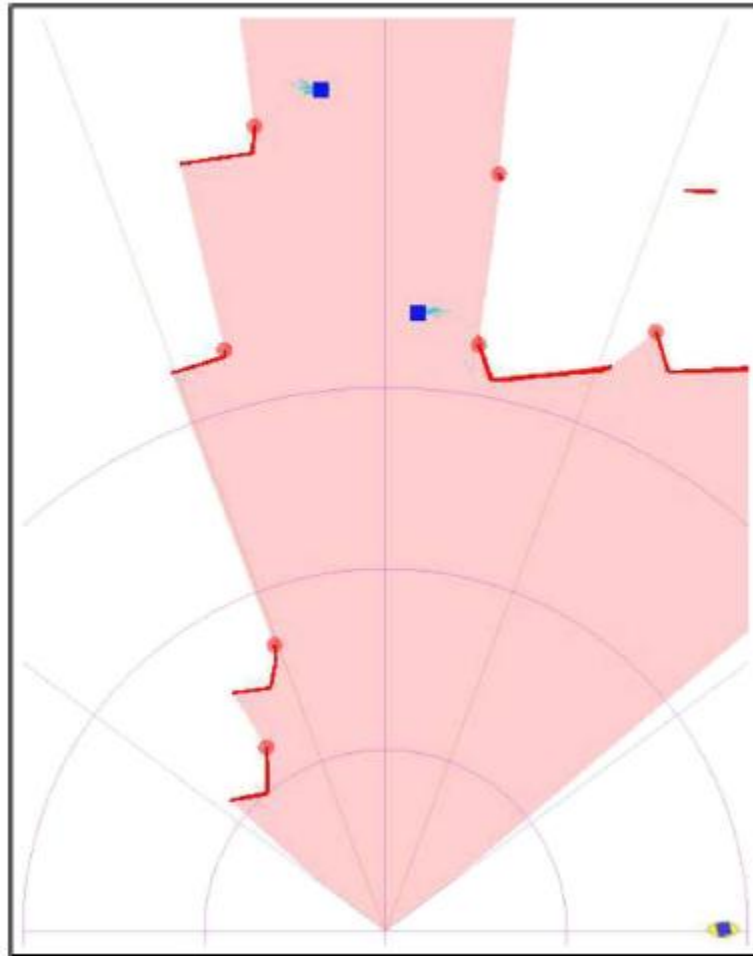


Ilustración 45. Áreas peligrosas y posibles peatones. (Círculos) Áreas peligrosas. (Cuadrados) Posibles peatones.

3.3.3 Fusión de datos

3.3.3.1 Clasificación

Una vez obtenidos los datos del láser y de la cámara de visión se procede a fusionarlos. La fusión es un desarrollo impulsado por el láser: Se utiliza un láser para generar un candidato, y la visión se utiliza para validar el candidato.

Lo primero que se realizará es una clasificación. Una vez que se encuentran las áreas de atención se activa la búsqueda de peatones en dichas áreas, y para ello se

utiliza la técnica de AdaBoost [9]. En lugar de utilizar sólo dos clases (peatones y no peatones), utilizaremos tres clases:

- 1) peatones;
- 2) no peatones;
- 3) peatones que aparecen.

Los peatones que aparecen son los peatones que no están inicialmente completamente visibles, es decir, parcialmente ocultos por los obstáculos, de manera que sólo puede ser enmarcada una parte de la forma de los peatones.

3.3.3.2 Alerta

Una vez que se detecta un peatón oculto en parte por un vehículo por el sistema de visión, el sistema emite una alerta interna, en este caso, sin avisar al conductor, ya que el nivel de peligro todavía no se ha determinado. Si el peatón entra en la zona de conducción entonces se le avisa al conductor, si se estima una posible colisión se activa automáticamente el freno del coche para evitarla. Es de fundamental importancia tener en cuenta que el conductor es advertido sólo cuando el peatón está completamente visible, no obstante el seguimiento por parte de los sensores se realiza una vez han detectado una parte del peatón.

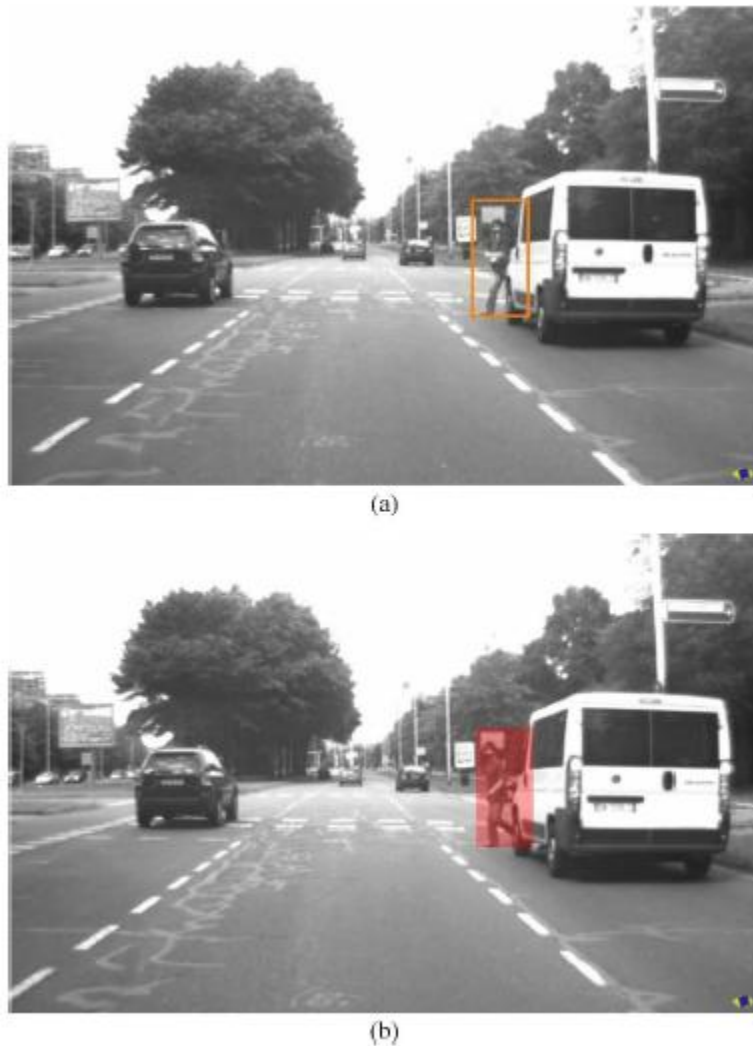


Ilustración 46. Dos *frames* de una secuencia. (a) El sistema detecta el peatón en la parte ocluida y emite una alerta interna, pero sin previo aviso al conductor. (b) Cuando el peatón detectado es totalmente visible sigue la alerta interna, y se envía un mensaje de advertencia de inmediato al conductor.

3.3.4 Resultados

Por lo general, la evaluación del desempeño de los sistemas de detección de peatones proporciona correctamente el porcentaje de peatones localizados y la tasa de falsas alarmas. Como este sistema se desarrolla bajo una nueva perspectiva, las cifras de rendimiento se ofrecen como el porcentaje de peatones que aparece identificado correctamente y la tasa de falsas alarmas. Se incluyeron diversos escenarios: caminos grandes y estrechos, aparcamientos, autopistas y caminos

rurales. Durante las pruebas por la noche, se desarrollaron escenarios con y sin iluminación externa.

La Tabla 1 muestra los resultados de la pruebas.

Tabla 1. Resultados
RESULTS OF A 10-H DRIVE (~236 km, ~540 000 IMAGES)

| Perform. index | Correct Detections | False Negatives | False Positives |
|----------------|--------------------|-----------------|-----------------|
| A | 24 | 11 | 1 |
| B | 5 | 0 | 0 |
| C | 8 | 0 | 0 |

En las pruebas, los índices de desempeño que se tuvieron en cuenta fueron:

- A) El número de peatones que apareció de repente frente al vehículo (que debe generar las advertencias para el conductor);
- B) El número de peatones que apareció en el frente del vehículo que han seguido caminando y que pueden ser golpeados por el vehículo (que debe dar lugar al freno automático);
- C) El número de peatones totalmente visible, que pueden ser atropellados por el vehículo (que debe activar el freno automático).

Para cada índice de rendimiento, se ha calculado el número de detecciones correctas, falsos positivos y falsos negativos. Un total de 24 peatones que aparecieron de repente fueron correctamente detectados en las pruebas (caso 1). Sólo está presente un falso positivo.

Todos los peatones fueron localizados (sin falsos negativos). No se detectaron falsos positivos debido a la especial atención que se presta para desarrollar este enfoque SDS. De todos modos, es importante tener en cuenta que, por razones de seguridad, dado que las pruebas fueron realizadas en tráfico real, el tamaño de la zona de desaceleración aumentó, sin embargo, el sistema se comportó de forma muy

satisfactoria.

El análisis de los resultados obtenidos durante la prueba fue satisfactorio y además también destacó los buenos resultados en el caso de lluvia, cuando los peatones con paraguas se han detectado también. De todos modos, el sistema no es capaz de discriminar entre varios peatones juntos en movimiento o en situaciones en las que el escáner láser no es capaz de obtener una imagen clara, por ejemplo, cuando los peatones tienen bolsas u otros objetos grandes. Por otra parte surgen problemas si de repente aparecen los peatones caminando muy despacio o peatones en condiciones de iluminación muy críticas.

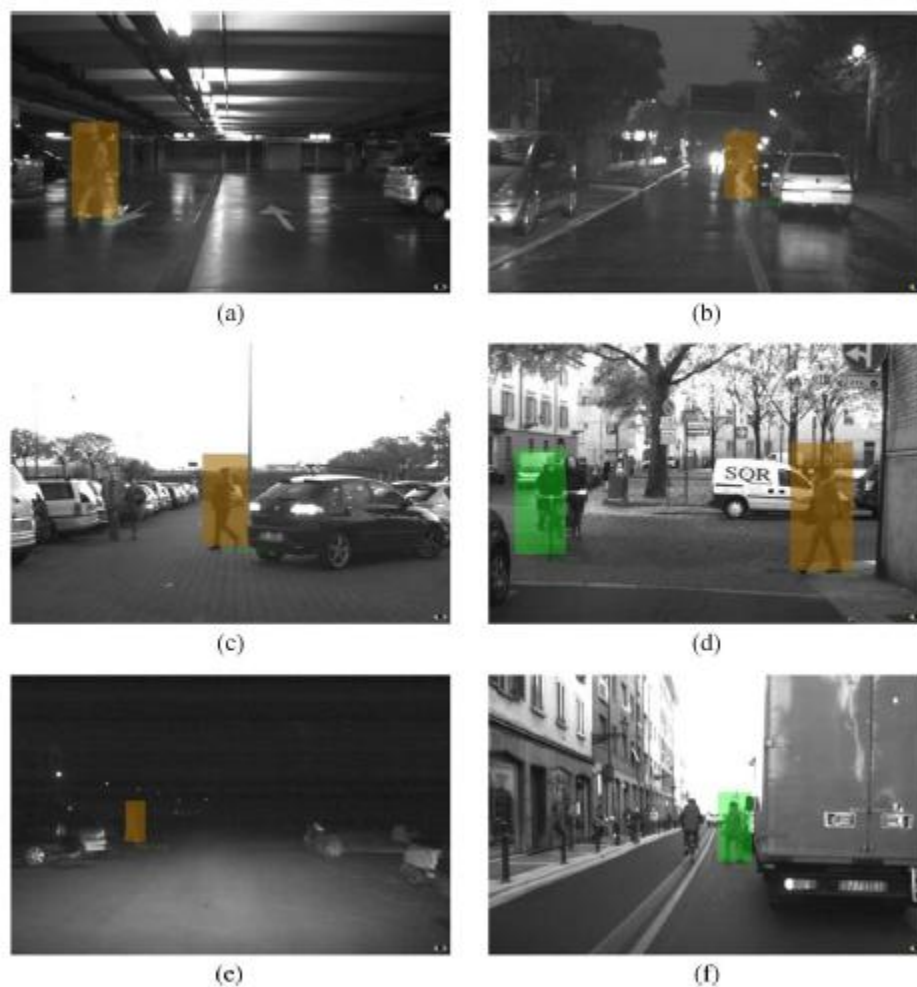


Ilustración 47. Se muestran algunos peatones que aparecieron de repente y fueron detectados correctamente (a) en un aparcamiento subterráneo, (b) bajo la lluvia, (c) detrás de un vehículo mal alineado, (d) tras una pared, y (e) en la noche, y (f) a un peatón que apareció de repente detectado como un peatón no peligroso (falso negativo).

3.4 Camera and Imaging Radar Feature Level Sensorfusion for Night

Vision Pedestrian Recognition [15]

Autores: *Matthias Serfling, Otto Loehlein, Roland Schweiger. Daimler AG, Group Research & Advanced Engineering; y Klaus Dietmayert. University of Ulm, Inst. of Measurement, Control and Microtechnology.*

3.4.1 Introducción

El enfoque presentado en este apartado corresponde a un sistema de fusión de un sensor de radar ACC (*Adaptive Cruise Control*) y una cámara de visión de infrarrojo cercano (NIR) con el fin de detectar a los peatones por la noche a una distancia de 120 metros en entornos rurales y suburbanos.

3.4.2 Sensores

La colocación de los sensores utilizados se muestra en la siguiente ilustración:



Ilustración 48. (a) Investigación de vehículos con sensor de posiciones de montaje y (b) generación avanzada del sensor radar ACC ARS300 de Continental AG.

La cámara va situada detrás del parabrisas y el sensor de radar en la parte delantera de la parrilla del radiador.

Con el fin de asociar las mediciones de ambos sensores, la alineación de cada sensor en el sistema del vehículo debe ser conocida con exactitud. El sistema de radar ha sido calibrado usando el procedimiento de alineación desarrollado por Continental AG por el que se alinea centro del haz del radar en *azimut* y en elevación con respecto al eje longitudinal del vehículo. Sin embargo, para calibrar el sistema de cámara se utiliza la matriz fundamental $\underline{F} = \underline{F}_{intr}\underline{F}_{extr}$ que contiene los parámetros intrínsecos. Los parámetros extrínsecos (posición de montaje y los ángulos de Euler) se evalúan mediante la resolución de un ajuste de datos no lineal por mínimos cuadrados:

$$\hat{\underline{F}}_{extr} = \arg \min_{\underline{F}_{extr}} \sum_{i=1}^N \left(\underline{F}_{intr} \underline{F}_{extr} \underline{\mathbf{x}}(r_i, \varphi_i) - \begin{pmatrix} u \\ v \end{pmatrix}_i \right)^2$$

Ecuación 6

3.4.3 Sistema de fusion

Con el fin de aumentar la robustez, ya que se requiere una sólida detección de peatones por la noche, el sistema combina la información de sensores complementarios en un mismo nivel de referencia, utilizando el algoritmo de entrenamiento AdaBoost para clasificar los posibles peatones.

AdaBoost selecciona gradualmente esas características con el menor error ponderado, hasta que el algoritmo alcanza una tasa de detección óptima de lo correcto y lo falso. Esta técnica asegura una selección y ponderación óptima de las características de los dos sensores para la tarea de clasificación. Sin embargo la medida del radar define una región de interés mucho mayor que el objeto medido. Además, la superficie del objeto no puede ser medida por el sistema del radar. De esta forma puede haber múltiples ventanas de búsqueda en la imagen que estén asociadas a una única medición del radar.

A fin de realizar un sistema de detección en tiempo real este sistema de fusión incorpora clasificadores en cascada.

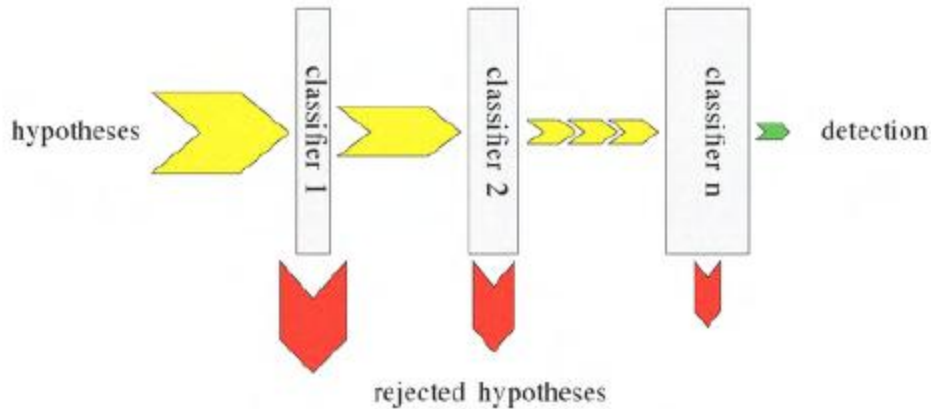


Ilustración 49. Clasificador en cascada: Cada etapa descarta hipótesis y pasa a la siguiente llevando muestras a la etapa sucesivas y más complejas.

Al aumentar el número de etapas, la complejidad de estos clasificadores aumenta por el número de características utilizadas. Cada etapa del clasificador se entrena con el algoritmo AdaBoost para elegir los rasgos más significativos de ambos sensores. Dependiendo de la capacidad de discriminación de las características desde el sensor de la cámara y el sensor de radar, la cascada resultante se puede realizar de manera muy eficiente.

En el plano de la imagen los peatones se representan por una ventana de búsqueda rectangular definida según la proporción entre altura y anchura.

En el plano de radar el conjunto de datos se compone de dos diferentes subconjuntos de filtros:

- Un subconjunto formado por los filtros que describen el valor medio de la intensidad del objeto reflejado.
- Un subconjunto que consta de filtros de señales complejas con una estructura similar a los filtros de radar CFAR (*constant false alarm rate*).

Estos filtros no son capaces de clasificar de forma inequívoca a un peatón. Sin embargo, se pueden utilizar para rechazar las muestras del fondo de manera muy eficiente.

El clasificador presentado en cascada requiere una tupla de la búsqueda de ventanas que indiquen el mismo escenario en el plano de la imagen y en el plano del radar para poder combinarlos entre sí. Cada tarea de clasificación requiere una constante evaluación del clasificador.

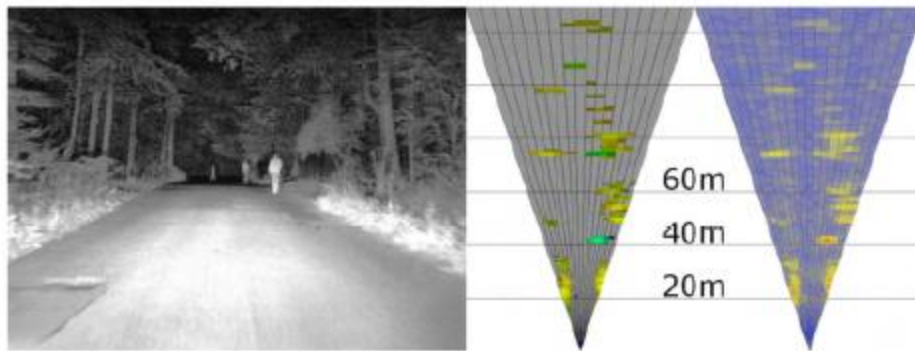


Ilustración 50. Las mediciones de radar: la imagen de la izquierda representa la escena tomada por la cámara. La imagen centrada muestra la intensidad de un umbral de radar (valor) y la velocidad relativa (tono). Los tres peatones en movimiento pueden ser reconocidos por el componente verde del color alrededor de 41m, 75m y 108m. La imagen derecha muestra todos los valores de intensidad de radar (el azul corresponde baja intensidad y el rojo a una mayor intensidad).

La ilustración 50 muestra una escena típica de los peatones en el medio rural. El etiquetado de los peatones en la imagen es bastante simple. Sin embargo, el etiquetado de peatones en la imagen del radar es mucho más complicado y requiere de un experto o algún tipo de asistencia de la etiqueta de la siguiente manera:

- 1) Estimación de la distancia monocular asumiendo un tamaño fijo de la etiqueta del peatón en el plano de la imagen.
- 2) Detectores CFAR aplicados a la línea epipolar proyectada en el plano de radar dada por un tamaño mínimo y máximo de los peatones.
- 3) Estimación de la distancia estéreo mediante el análisis de la disparidad de píxeles dentro de la etiqueta en el plano de la imagen.

3.4.4 Resultados

Una vez probado se puede ver que el detector de fusión supera todas las expectativas. El resultado en cascada supera significativamente el detector de referencia y además requiere menos etapas y por lo tanto menos características (ver Tabla 2).

Tabla 2. Características del detector.

| | stages | image features | radar features |
|-----------|--------|----------------|----------------|
| FUSION | 28 | 1061 | 407 |
| REFERENCE | 30 | 1621 | - |

Además AdaBoost seleccionó hasta un 38% de las características del conjunto de características de radar que describe su poder discriminativo. Estas características de radar permiten una rápida reducción de las hipótesis en las primeras etapas de la cascada detector de fusión. La ilustración 51 compara el cálculo de los clasificadores en entornos de diferente complejidad.

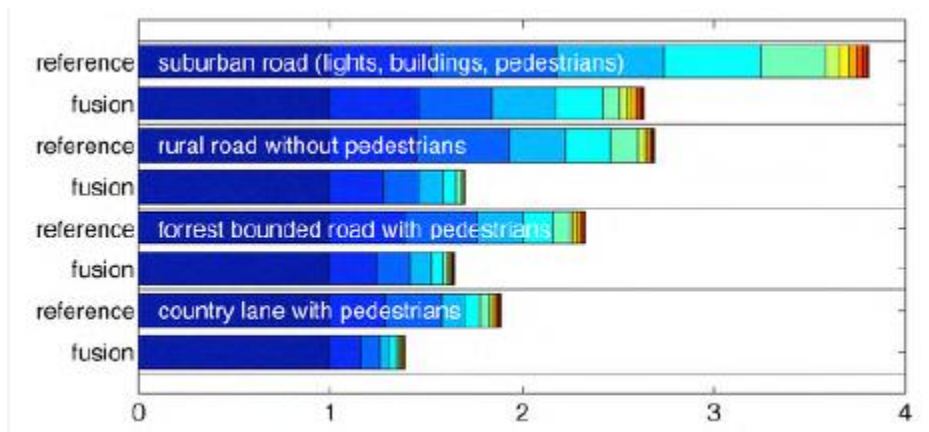


Ilustración 51. Análisis de costo computacional en las primeras etapas de la cascada hasta la capa 12: El detector de imagen de referencia basado en la fusión y el detector clasifican el mismo conjunto de hipótesis. El detector de fusión muestra resultados también similares en ambientes altamente estructurados tales como las zonas suburbanas, con edificios y en gran medida los cambios de condiciones de luminosidad.

La ilustración 50 muestra las curvas ROC para comparar el desempeño del detector de fusión y el detector basado en la imagen de referencia. La tasa de detección y falsa alarma es mucho mejor en el sistema de fusión que en el sistema de referencia utilizando sólo la información proporcionada por la cámara de infrarrojo cercano. El detector de fusión utiliza la información proporcionada por el sensor de radar para descartar hipótesis válidas, el clasificador de referencia no puede distinguir entre los peatones hipotéticos en diferentes distancias. Por lo tanto, esta última produce resultados de la detección incapaz de localizar la zona peatonal en el espacio:

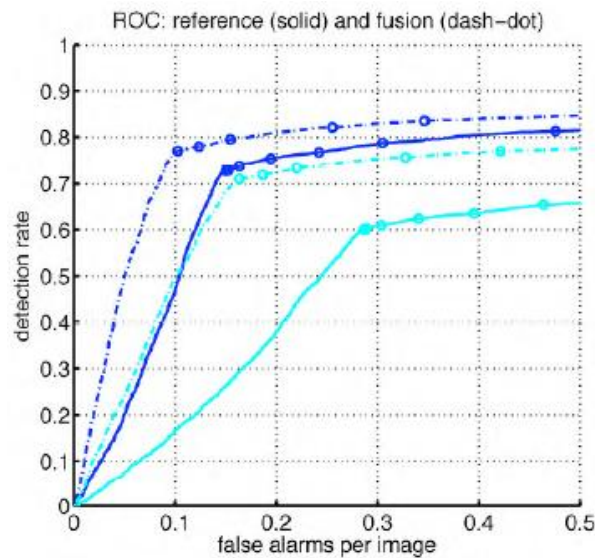


Ilustración 52. Comparación de rendimiento entre el detector basado en la imagen de referencia detector (línea continua) y el detector de fusión (línea discontinua) para dos coberturas 0,5 (azul, los dos gráficos superiores) y 0,6 (cian, gráficos inferiores).

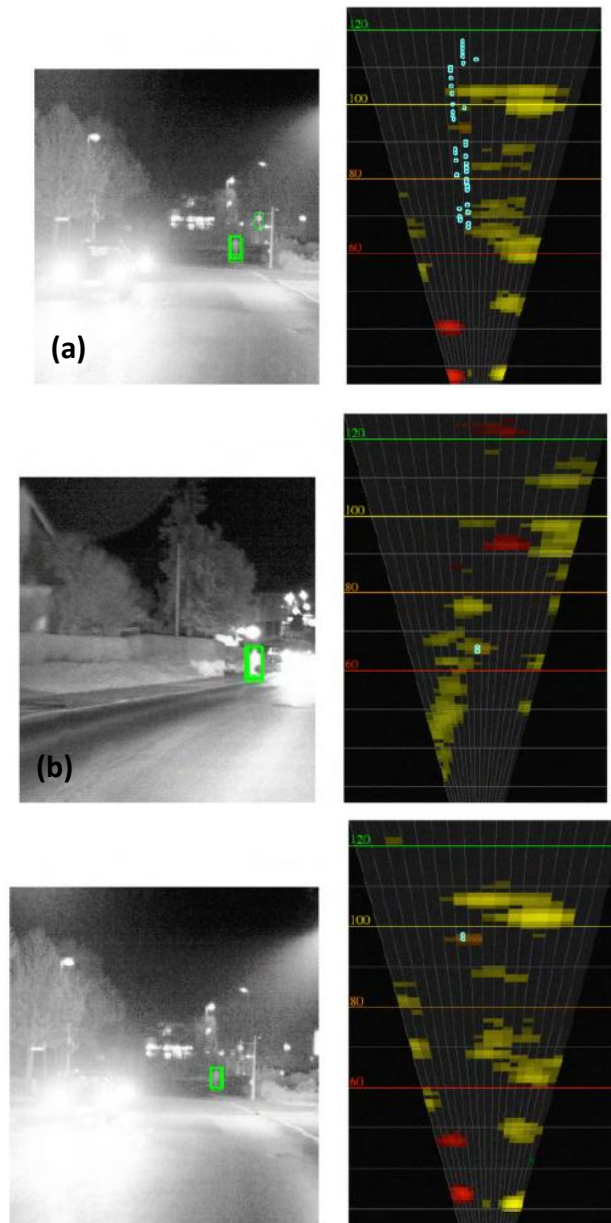


Ilustración 53. Se pueden ver ejemplos de detección en un entorno suburbano (imagen de la cámara con detecciones verde a la izquierda y los datos de radar con detecciones cian derecha): (a) Típico resultado del detector basado en la imagen de referencia. La incertidumbre en el espacio de localización es muy alta con características sólo para la clasificación de la imagen. Las figuras (b) y (c) muestran el mayor rendimiento del detector de fusión que incorpora también las características proporcionadas por el sistema de radar. El resultado es una localización precisa de los peatones en (b) 65m y 100m (c).

En este enfoque de fusión de dos sensores, el radar y la cámara de visión pueden rechazar la hipótesis en las primeras etapas de la cascada y ofrece una localización espacial precisa de las hipótesis del objeto resultante.

3.5 Pedestrian detection based on maximally stable extremal regions [17]

Autores: Vadim Frolov y Fernando Puente León. *Institute of Industrial Information Technology, Karlsruhe Institute of Technology.*

3.5.1 Introducción

A continuación se presenta un enfoque de detección de peatones basado en *regiones extremas de máxima estabilidad* (MSER) mediante la fusión de un sensor lidar y una cámara infrarroja. Estas regiones se extraen a partir de imágenes de infrarrojos y se utilizan como hipótesis sobre la presencia de peatones y su ubicación. Con el fin de extraer las dimensiones apropiadas de las cajas alrededor de un potencial peatón, el algoritmo realiza una fusión de los datos de un sensor lidar con la información proporcionada por las MSER. La lista resultante de *cajas* se filtra mediante una medida de dispersión para descartar otros objetos no humanos.

3.5.2 Sensores

En cuanto a los sensores utilizados, la cámara de infrarrojos percibe el entorno y proporciona información sobre la forma y la temperatura. Estos datos son preprocesados para extraer las regiones cálidas. Las áreas cálidas resultantes son etiquetadas. Cada una de las áreas cálidas extraídas se considera como una potencial cabeza humana. El escáner lidar se utiliza para realizar un análisis de una línea de la escena con una resolución angular $\Delta\phi$ de 0.25 a 1° y un ángulo de apertura de 180°. La escena está revisada cada Δt segundos.

Los dos sensores están sincronizados ya que la velocidad de la cámara y el lidar se han definido para ser iguales. Ahora, las señales deben estar registradas con el fin de asociar a cada zona cálida de la señal IR con un objeto detectado por el lidar. Para obtener una correlación entre los datos infrarrojos y el lidar, ambos sensores se alinean verticalmente con la cámara sobre el sensor lidar. De esta manera, la

proyección de escaneado del sensor lidar coincide con la proyección horizontal de la cámara.

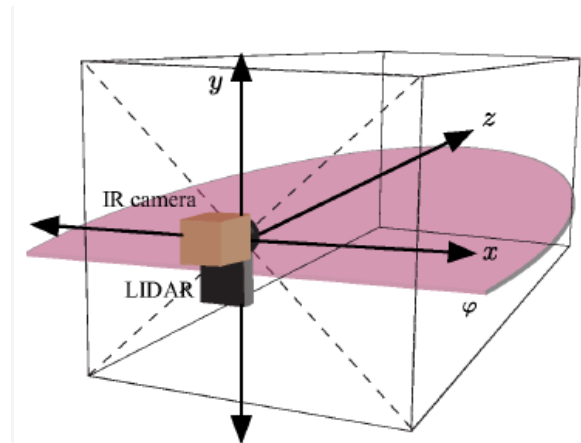


Ilustración 54. Configuración de la plataforma del sensor

3.5.3 Hipótesis sobre la presencia de peatones

3.5.3.1 Detección de áreas cálidas

A continuación se explican las hipótesis sobre la presencia de peatones. La imagen infrarroja de un objeto depende de la cantidad de calor que este emite: cuanto más caliente está el objeto, más brillante parece. En la mayoría de los casos la temperatura de un peatón es más alta que la de su entorno. La primera fase del algoritmo que se va a utilizar centra la atención en las áreas de la imagen de entrada con un valor de alta intensidad, que representan los objetos cálidos. Esto se consigue utilizando las *regiones extremas de máxima estabilidad o MSER*, que están conectados los componentes de una imagen apropiada.

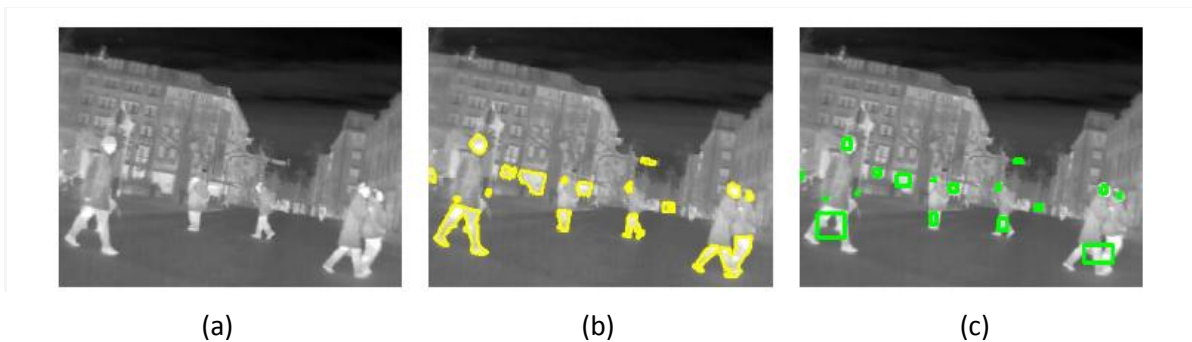


Ilustración 55. Detección de áreas cálidas. (a) Imagen original. (b) Curvas de nivel sobre la base de la MSER. (c) Formas rectangulares, como consecuencia del enfoque MSER.

La detección de las MSER está relacionada con un umbral, ya que cada región extrema es un componente conectado a un umbral de la imagen. Sin embargo, no se busca ningún umbral óptimo o global, todos los umbrales son probados y se evalúa la estabilidad de los componentes conectados. La salida del detector de MSER no es una imagen binaria. En algunas partes de la imagen, existen varios umbrales estables y en este caso se emite un sistema de subconjuntos anidados. Se considera que cada *caja* de unión contiene potencialmente una cabeza humana.

3.5.3.2 Extracción de las cabezas de la imagen

Para llevar a cabo una clasificación completa con éxito los hipotéticos peatones deben ser extraídos de la imagen. Luego, para cada posible cabeza humana, se define una región a escala, que puede contener el cuerpo completo. Cada región se refiere como región de interés del candidato (candidato ROI).

3.5.3.3 Definición de las regiones de interés

A continuación los candidatos ROI situados cerca de la misma posición y y que forman una intersección entre sí se agrupan en una *caja* de delimitación más grande. Esto permite excluir dimensiones de las *cajas* que corresponden a partes del cuerpo como los brazos o las piernas de otros cálculos. Además se rechazan las regiones que no contienen peatones mediante una medida de dispersión. En la

ilustración 56 se puede observar un ejemplo de la medida de dispersión (*dispersedness*):

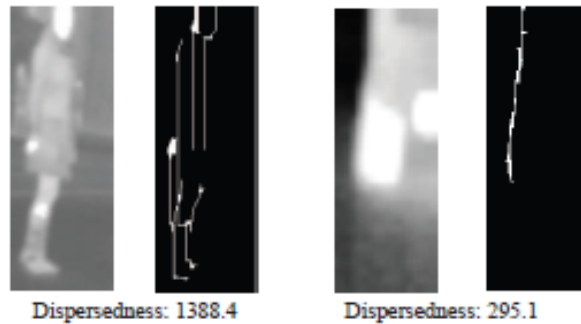


Ilustración 56. Valores típicos de la medida de dispersión y desorden.

En las escenas reales de tráfico, los peatones deben estar separados de partes de los automóviles, partes de los edificios y una mezcla de piezas de varios objetos. En este caso, la medida de dispersión de un peatón es mayor que el de otro objeto. La medida de dispersión es una suma de las medidas de dispersión de todas las regiones en una imagen resultante. Después, se aplica a todas las medidas un umbral de valor basado en los datos experimentales. Las regiones en las que el valor de la medida de dispersión es menor que el valor de umbral se rechazan. Todas las demás regiones se refieren como regiones de interés. En la ilustración 57 se muestra un ejemplo del proceso de filtrado basado en la medida de dispersión de una escena, las dos primeras contienen peatones y las dos últimas no:

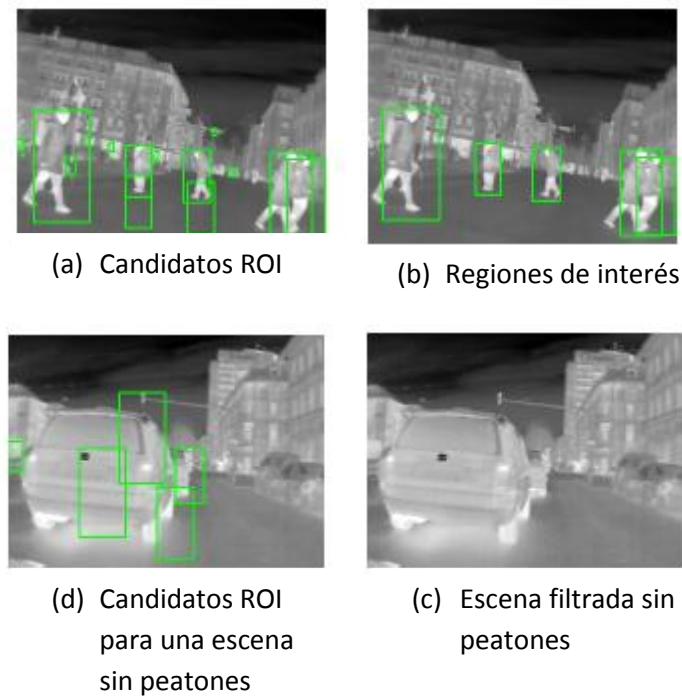


Ilustración 57. Proceso de los resultados de MSER

3.5.4 Resultados

El sistema desarrollado ha sido probado con secuencias pregrabadas de imágenes de infrarrojos y señales correspondientes del lidar. Las medidas han sido tomadas desde un vehículo en movimiento. El sistema procesa cada fotograma por separado. La secuencia resultante de etiquetado de datos por infrarrojos consistió en 1.000 frames. El número de regiones extraídas varía de acuerdo con el umbral que utiliza para encontrar MSERs. La disminución exponencial del número de las regiones extraídas se muestra en la ilustración 49 (a). A partir del valor de umbral 11, el número de objetos extraídos es de alrededor de 40 por cuadro y, como se muestra en la ilustración 58 (b), los peatones son extraídos. Por lo tanto, el valor de umbral 11 se utiliza para un mayor procesamiento.

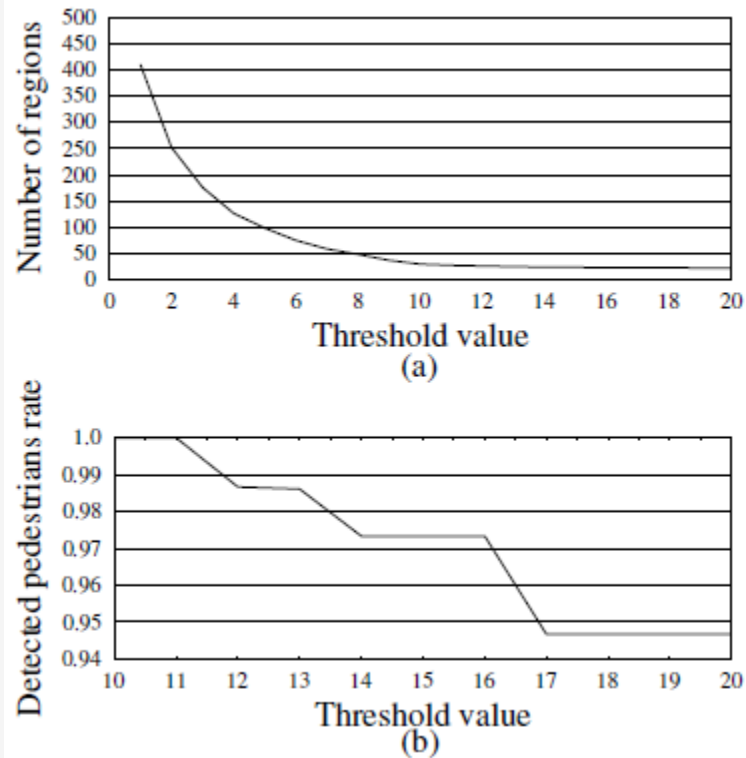


Ilustración 58. Extracción de MSER en base al umbral: (a) número de regiones extraídas, (b) tasa de regiones extraídas que contienen peatones.

Los objetos con un tamaño similar y la forma de representar a un peatón son la causa más frecuente de detecciones erróneas. Esto significa que se debe desarrollar un proceso de validación correcto. Un proceso de validación eficaz no puede basarse en una única medida de dispersión. Para evaluar el rendimiento del enfoque propuesto, se tiene que realizar una búsqueda con parámetros que coincidan con el enfoque de dos modelos diferentes que detectan la cabeza del peatón. El primer modelo codifica una máscara binaria que muestra una cabeza blanca sobre un fondo negro. Y el segundo modelo codifica la forma de la cabeza. La primera curva ROC que se muestra en la ilustración 60 (a) se ha obtenido variando el umbral de la medida de dispersión utilizada para la detección.

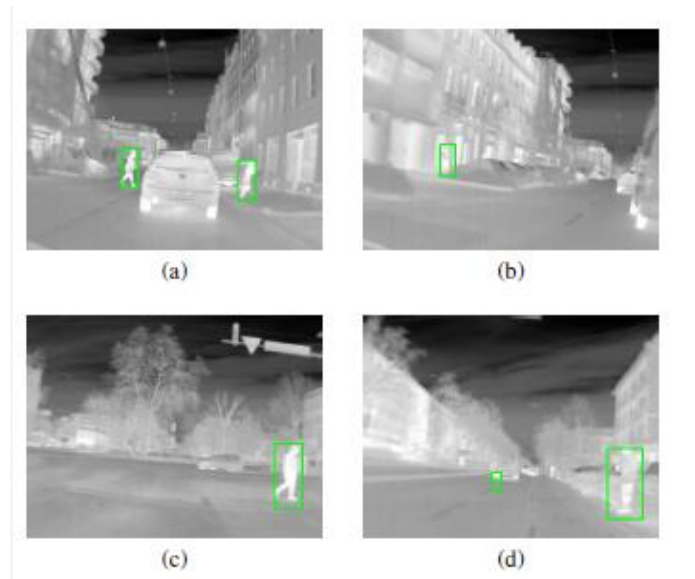


Ilustración 59. Resultados de la detección: (d) muestra un falso positivo.

Se puede observar que en esta condición, el sistema es capaz de detectar correctamente el 90% de los peatones, casi sin falsas detecciones. La ilustración 50 muestra algunos resultados de este tratamiento: las formas humanas detectadas están marcadas con un cuadro delimitador de color verde. La segunda curva ROC (ilustración 41(b)) se ha obtenido variando el umbral de correlación utilizado en la localización del enfoque de la cabeza. En este caso el sistema es capaz de detectar correctamente más del 80% de los peatones con una baja tasa de falsas detecciones.

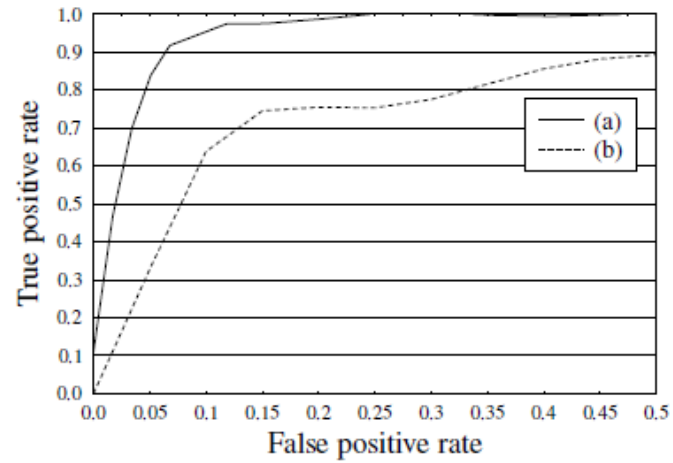


Ilustración 60. Curvas ROC para los resultados basados en (a) medida de dispersión y (b) plantilla de unión.

3.6 Non-parametric Laser and Video Data Fusion: Application to Pedestrian Detection in Urban Environment [18]

Autores: *S. Gidel, C. Blanc, T. Chateau, P. Checchin y L. Trassoudaine. LASMEA - UMR 6602 CNRS Blaise Pascal University. Aubière, France*

3.6.1 Introducción

Este método se lleva a cabo en el marco del Proyecto LOVE (*Software for vulnerable observation*), que tiene como objetivo mejorar la seguridad vial, centrándose principalmente en la seguridad de los peatones. Para resolver el problema de la detección y la identificación es lógico utilizar sensores de vídeo en un primer momento, dada la capacidad de este tipo de sensor para detectar y analizar el tamaño, la forma y la textura de un peatón. Sin embargo en muchas circunstancias adversas resulta una tarea difícil. Por otro lado, la capacidad de un sistema láser para detección de peatones ha sido demostrada con éxito anteriormente para detectar y realizar el seguimiento del peatón, sin embargo, las evidentes limitaciones de este sensor, su sensibilidad a las condiciones atmosféricas como la lluvia y la niebla y las oclusiones frecuentes entre los objetos, requiere diseñar un método de fusión de láser y cámara de visión para mejorar un sistema que pueda evitar un accidente de tráfico de peatones. Por lo tanto, dos problemas distintos tienen que ser resueltos en forma conjunta: los datos de la asociación y la estimación. En este método, se presentará una fusión de un nuevo sistema de láser / cámara, que tiene como objetivo la fiabilidad, control en tiempo real y el seguimiento de varias personas en un entorno urbano.

3.6.2 Sensores

En cuanto a los sensores utilizados, éstos son un sensor láser de Ibeo y dos sensores de visión del equipo de pruebas de Renault, montados como se indica en la ilustración 61:

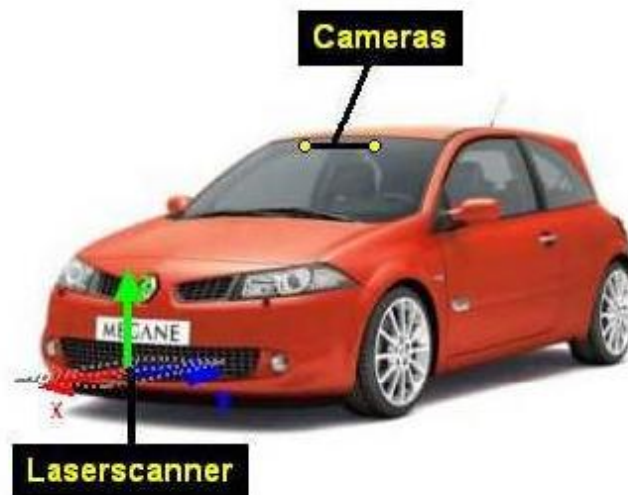


Ilustración 61. Localización de los sensores en el vehículo de prueba

3.6.3 Fusión de los datos de los sensores

3.6.3.1 Arquitectura del sistema

La arquitectura del sistema de detección de peatones propuesto se muestra a continuación en la ilustración 62.

La detección de peatones se realiza mediante el escáner láser y los datos proporcionados por la cámara de visión. Después se desarrolla un módulo de fusión centralizada cuyas principales aportaciones son:

- Asociación no-paramétrica de datos,
- Mezcla de distribuciones gaussianas y uniformes para probabilidad de computación,
- Cálculo de un factor de confianza de fusión.

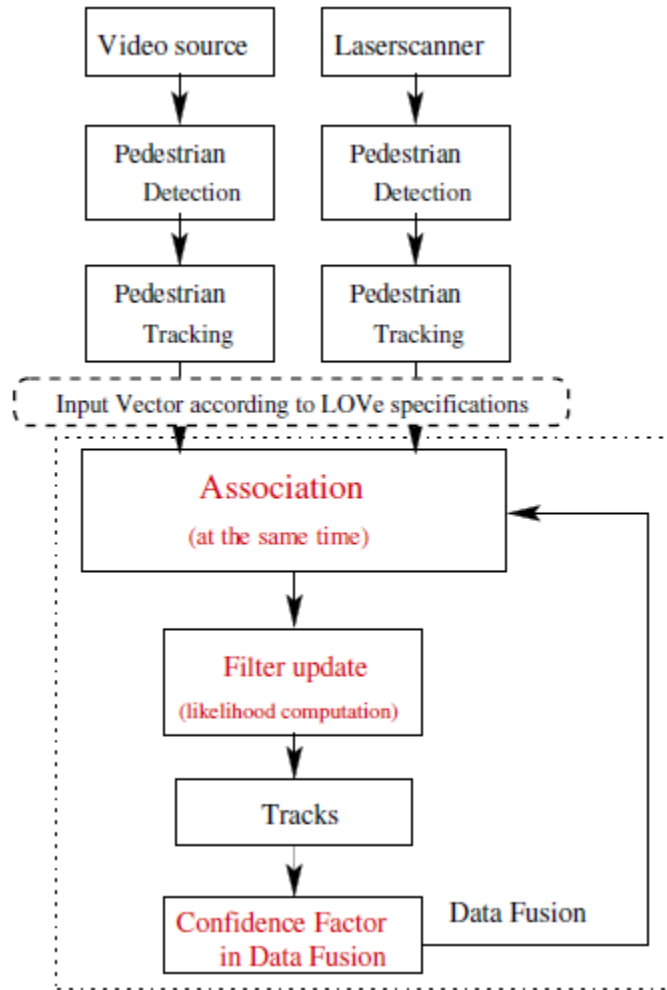


Ilustración 62. Arquitectura multi-módulo con información de lidar y visión para la detección de peatones y su clasificación. Destacan en color rojo las principales aportaciones.

3.6.3.2 Asociación no-paramétrica de datos

Para asociar los datos utilizamos un método no-paramétrico. Los métodos no-paramétricos permiten tomar en cuenta las muestras y su distribución espacial en los parámetros del espacio. N_x es el número de objetivos para realizar un seguimiento. Este número no se conoce en el instante k . El seguimiento consiste en la estimación del vector de estado obtenido mediante la concatenación del vector de N_x de todos los objetivos. El vector $X_k = \{x_k^{1,l}, \dots, x_k^{N_x,l}\}_{l=1}^N$ viene dada por la ecuación de estado $X_k = f(X_{k-1}) + W_k$ descompuesta en las ecuaciones de N_x :

$$\mathbf{X}_k = F_k(\mathbf{X}_{k-1}, \mathbf{W}_k)$$

Ecuación 7

Donde \mathbf{X}_k representa el vector de estado en el instante k ; N es el número de partículas, y los ruidos \mathbf{W}_k se suponen que son espaciales y temporalmente blancos. El vector de observación recogido en el tiempo k se denota por $\mathbf{Z}_k = (z_k^1, \dots, x_k^{N_z})$ siendo N_z el número de observaciones deducidas del proceso.

La matriz de asociación A_k se introdujo para describir la asociación entre las mediciones y los objetivos $\mathbf{Z}_k \mathbf{X}_k$. Se ha elegido un marco no paramétrico con el fin de estimar la asociación de la matriz A_k . Para hacer posible la generación de una sucesión de las áreas que responden a las buenas condiciones de la estimación utilizaremos el método de “la Ventana de Parzen”:

Este método no está supervisado, así que no hay conocimiento previo para el proceso de fusión de datos. La función de verosimilitud $p(z_k^j, x_k^i | w_h)$ representa la probabilidad de que una partícula en 2D pertenezca a los datos fusionados. La probabilidad $p(z_k^j, x_k^i | w_h)$ es el modelo de una ventana Parzen que calcula la distancia entre una observación z_k^j ubicada en la imagen y todos sus vecinos $x_k^{i,l}$, tales como:

$$p(\mathbf{z}_k^j, \mathbf{x}_k^i | w_h) = \frac{1}{N} \sum_{l=1}^N \varphi(\mathbf{z}_k^j, \mathbf{x}_k^{i,l})$$

Ecuación 8

Donde $\varphi(z_k^j, x_k^{i,l})$ es la función del núcleo que permite modificar la zona de influencia de la observación con sus vecinos. Se utiliza una mezcla de distribuciones gaussianas y uniformes para fusionar en una misma distribución toda la información disponible de las salidas del algoritmo de un solo sensor. $\varphi(z_k^j, x_k^{i,l})$ viene dada por la siguiente ecuación:

$$\varphi(\mathbf{z}_k^j, \mathbf{x}_k^{i,l}) = (1 - \gamma_k^j) \cdot \mathcal{U}\left(\frac{-1}{\gamma_k^j}, \frac{1}{\gamma_k^j}\right) + \gamma_k^j \cdot \exp[-\lambda_c \cdot d_c(\mathbf{z}_k^j, \mathbf{x}_k^{i,l})]$$

Ecuación 9

El parámetro λ_c permite ajustar los valores. La distancia entre puntos d_c está definida por:

$$d_c(\mathbf{z}_k^j, \mathbf{x}_k^{i,l}) = (\mathbf{z}_k^j - \mathbf{x}_k^{i,l}) \Sigma_\varphi^{-1} (\mathbf{z}_k^j - \mathbf{x}_k^{i,l})^T$$

Ecuación 10

Con la matriz de covarianza Σ_φ , el resultado de la suma entre la matriz de covarianza es Σ_{SP2} dada por el algoritmo de seguimiento (que representa la variación en la posición de los peatones) y la medición del ruido de la matriz de covarianza R :

$$\Sigma_\varphi = \Sigma_{SP2} + R$$

Ecuación 11

La partícula de $x_k^{i,l}$ que tiene la mayor probabilidad es elegida por el estimador de máxima verosimilitud a partir de la siguiente ecuación:

$$A_k^{i,j}(i, j) = \arg \max_{(i,j)} (p(\mathbf{z}_k^j | \mathbf{x}_k^i) | w_1))$$

Ecuación 12

3.6.3.3 Cálculo del factor de la confianza en la fusión de datos

Todos los objetos rastreados se ponen a prueba para determinar si son o no el resultado de la fusión entre el láser y los datos de vídeo. Para ello, cada objetivo se evalúa mediante el cálculo de su Factor de Confianza en los Datos de Fusión (CFDF). Existen tres criterios que constituyen el CFDF: la confianza en la edad (tiempo)

de la pista (CAT), la tasa de detección global (DOR) y la tasa de fusión de sensores (SFR). El CAT permite evaluar si el objetivo de peatones se ha seguido durante mucho tiempo o no. El DOR, a condición de algoritmos de un solo sensor, es la tasa de confianza en que el objeto detectado es en realidad un ser humano. La SFR permite evaluar si el seguimiento es el resultado o no de los datos de la fusión entre el láser y las mediciones de la cámara de visión. El CAT y la SFR se calculan a partir de una distribución de Gauss:

$$CAT(t) = \begin{cases} \frac{1}{\sigma_o \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu_o}{\sigma_o}\right)^2\right] & 0 < t \leq \mu_o \\ 1 & t > \mu_o \end{cases}$$

Ecuación 13

Donde μ_o representa el tiempo mínimo de vida de una pista sin observación, σ_o permite reducir más y menos rápidamente el CAT y t es la edad de la pista.

$$SFR(x) = \frac{1}{\sigma_f \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_f}{\sigma_f}\right)^2\right]$$

Ecuación 14

Donde μ_f representa la relación teórica entre el número de los datos de láser y el número de datos de vídeo, σ_f permite la disminución más o menos rápidamente de la tasa de fusión de sensores y es x la relación entre el número de datos por láser y el número de datos de vídeo. Finalmente, el resultado final viene dado por:

$$CFDF = \frac{CAT + SFR + DOR}{3}$$

Ecuación 15

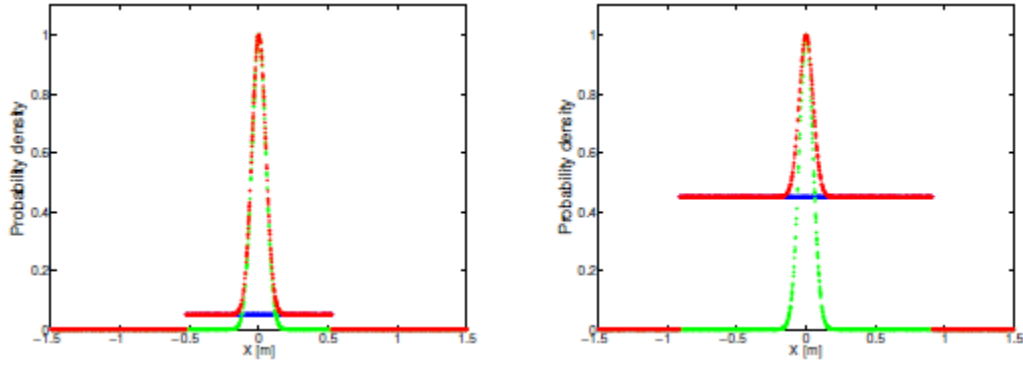


Ilustración 63. Un ejemplo de una mezcla de distribuciones de Gauss y uniforme, en azul la distribución uniforme, de color verde la distribución de Gauss y en rojo la mezcla de las distribuciones. A la izquierda, $s = 0,15$ m y $DOR = 0,95$. A la derecha, $s = 0,15$ m y $DOR = 0,55$.

A continuación se presenta el algoritmo de fusión de los datos del láser y la cámara de visión para mejorar el seguimiento de los peatones:

Algoritmo no paramétrico de Fusión de Datos

1. $K = 0$, generar N muestras de cada medición $j = 1, \dots, N_z$, i.e. $\{x_0^{j,l}\}_{l=1}^N = \{x_0^{1,l}, \dots, x_0^{N_z,l}\}_{l=1}^N$ donde $x_0^{j,l} = p(X_0^j)$

2. Calcular la matriz A_k para todas las mediciones y objetivos (N_z, N_x).

Si $(A_k \leq \alpha)$ con $\alpha =$ el umbral de la puerta:

$A_k^{i,j} = p(z_k^j, x_k^i | w_h)$ donde $p(z_k^j, x_k^i | w_h)$ es la probabilidad de asociación para la hipótesis i usando N partículas.

En caso contrario:

$A_k^{i,j} = 0$, entonces $\{x_k^{i,l}\} = \{x_{k-1}^{i,l}\}$ ir al paso 5.

3. Calcular los pesos $w_k^{i,l} = L_k^{i,j}$ y normalizar

4. Generar un nuevo conjunto $\{x_k^{i,l*}\}_{l=1}^N$ por muestreo para la sustitución n veces de $\{x_k^{i,l}\}_{l=1}^N$, donde $\Pr(x_k^{i,l*} = x_k^{i,l}) = w_k^{i,l}$

5. Predecir (simular) nuevas partículas usando diferentes ruidos realizados por las partículas. Calcular para cada estimación de su CFDF.

6. Aumento de k , y repetir el punto 2. Este punto es importante porque, cuando la clasificación para peatones no es fiable, las hipótesis (partículas) deben mantenerse en a fin de corregir un posible error en el estimador de estado.

3.6.4 Experimentos

En esta sección se presentan las simulaciones y los experimentos que permiten validar el algoritmo de fusión del láser y de datos de vídeo.

3.6.4.1 Simulaciones

El objetivo de estas simulaciones es mostrar cómo funciona el algoritmo de fusión de datos mejorando el sistema de seguimiento de peatones. En primer lugar, se propone un estudio del algoritmo de fusión. En la ilustración 64, se ha generado un escenario con dos peatones detectados a partir de datos láser y vídeo. En segundo lugar, se propone un estudio de cálculo de probabilidad. En la ilustración 65, se genera una nube de partículas al azar (puntos rojos). Las partículas (estrella azul) que representan el centro de la nube se seleccionaron como las mediciones. Con la misma incertidumbre acerca de la posición, los resultados son diferentes de acuerdo con la "tasa de detección en general". Este último punto es importante porque, cuando la clasificación para peatones no es fiable, las demás hipótesis (partículas) deben mantenerse en a fin de corregir un posible error en el estimador de estado.

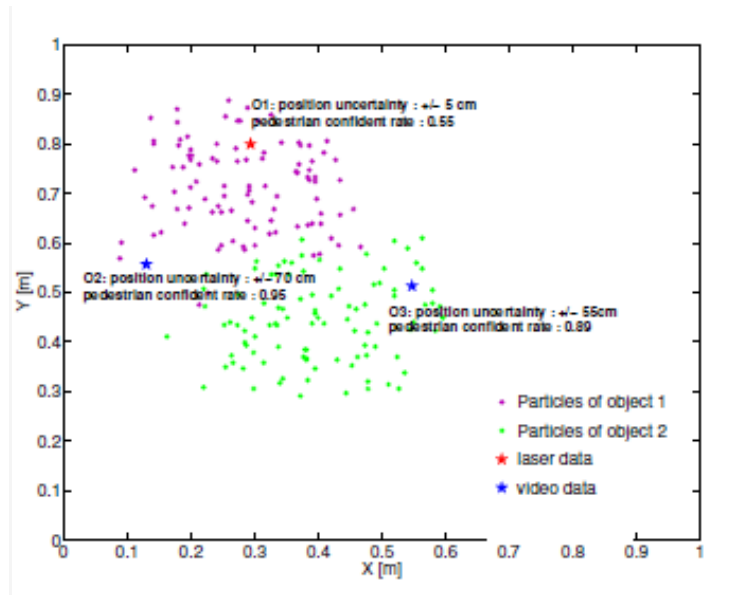


Ilustración 64. Asociación de datos de varios sensores. Aquí, de acuerdo con el criterio del vecino más cercano, O1 se asociaría a l objeto 1 y O3 al objeto 2, la asociación correcta está dada por el algoritmo de Parzen.

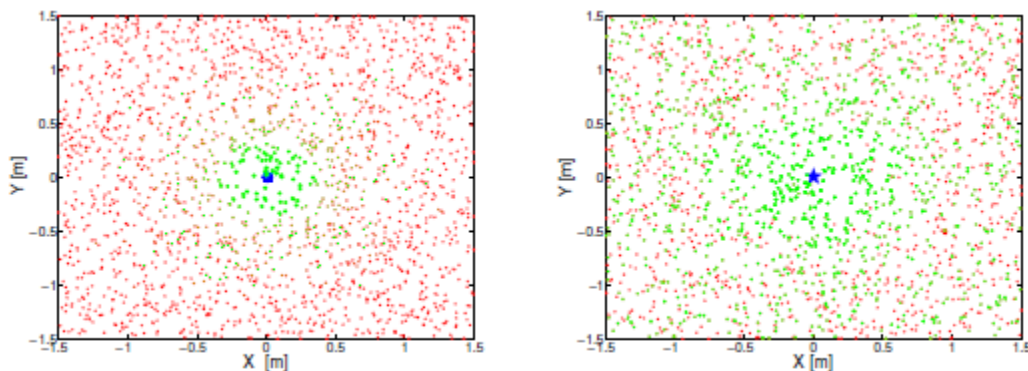


Ilustración 65. Ejemplo de cálculo de probabilidad en una nube de partículas en la ilustración 33. En la izquierda, $s = 0,15$ m y DOR = 0,95 y en la derecha, $s = 0,15$ m y DOR = 0,55.

3.6.4.2 Experimentos con datos reales

En los experimentos reales los peatones se mueven en todas las direcciones. El vehículo se mueve a una velocidad de 0 a 50 km/h, lo que permite poner a prueba la robustez de este método. Por lo general, en el marco de clasificación de peatones, las mediciones del lidar pueden generar pistas falsas. Estas mediciones del lidar dan como resultado en la mayoría de los casos de objetos adecuados fija para un ambiente urbano. Las falsas mediciones del lidar se pueden deber a objetos tales barreras de

seguridad, postes, árboles, etc. Para cada iteración, el número de falsas detecciones se obtiene mediante el cálculo de la relación:

$$\text{rate_of_false_detection} = \frac{N_T - N_P}{N_T}$$

Ecuación 16

con N_T el número total de detecciones y N_P el número de detecta los peatones. La tasa de detección de peatones (s) está dada por el cálculo de la relación:

$$\text{rate_of_pedestrian_detection} = \frac{N_P}{N_{P_VT}}$$

Ecuación 17

donde N_{P_VT} es el número de peatones que se encuentran actualmente en el área observada por los sensores. La Tabla 3 muestra la ventaja de fusión de datos para reducir significativamente el número de falsas detecciones cuando se utilizan un único lidar o una sola cámara. También se puede notar que la tasa de peatones detectados es más alta después de la fusión de datos.

Tabla 3. Tasa de detecciones falsas y correctas: cuando sólo se utiliza la cámara o el lidar, y después de la fusión de datos.

| | Lidar only | Camera only | After fusion |
|---------------------------|------------|-------------|--------------|
| False Detection Rate | 0.536 | 0.274 | 0.108 |
| Pedestrian Detection Rate | 0.916 | 0.702 | 0.928 |

En los experimentos realizados de puede comprobar el éxito de este método. A continuación se muestran algunas imágenes de la detección. Los puntos rojos representan el láser y los rectángulos azules la cámara de detección. Los rectángulos amarillos son el resultado de la fusión de ambos sensores.

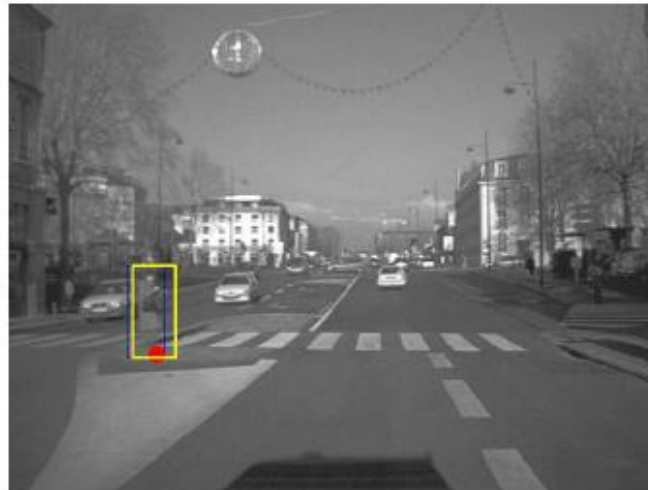


Ilustración 66. Ejemplo de detección en una sección transversal después de una fusión centralizada de datos de imágenes lidar y de vídeo. El punto rojo representa la detección lidar y los rectángulos azules representan cámara de detección. Los rectángulos amarillos son los resultados proporcionados por el módulo de fusión de datos.



Ilustración 67. Ejemplo de detección en un aparcamiento después de una fusión centralizada de datos de imágenes lidar y de vídeo. Se puede observar la correcta detección de un peatón a una distancia de hasta 25 metros.

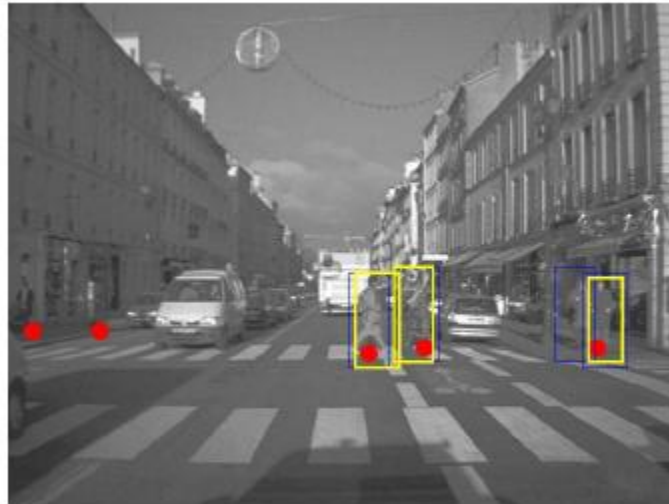


Ilustración 68. Ejemplo de detección en un aparcamiento después de una fusión centralizada de datos de imágenes lidar y de vídeo. Se puede observar la correcta detección de varios peatones en direcciones distintas.

3.7 An application of data fusion combining laser scanner and vision in real-time driving environment recognition system [13]

***Autores:** Chao-Hua Guan, Jian-Wei Gong, Yong-Dan Chen y Hui-Yan Chen. Intelligent Vehicle Research Center, Beijing Institute of Technology, Beijing, China.*

3.7.1 Introducción

En este método se llevará a cabo una calibración sencilla y rápida de una cámara de visión y un escáner láser combinado con una aplicación práctica del entorno de conducción en un sistema de reconocimiento en tiempo real, además del procesamiento de los datos del escáner láser para obtener una imagen de región de interés (ROI).

3.7.2 Estructura del sistema

En cuanto a los sensores utilizados, el sistema está equipado con un escáner láser montado en el paragolpes delantero y una cámara enfocada en la dirección de movimiento del vehículo. Como se muestra en la ilustración 69, las coordenadas del sistema se pueden definir así:

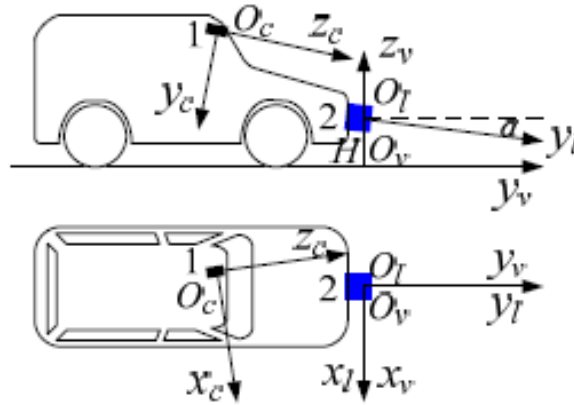


Ilustración 69. Descripción de la estructura del sistema y la posición de sensor. 1 es la cámara y el 2 es el escáner láser. α es el grado de inclinación del escáner láser. H es la altura de montaje del escáner láser.

- a) $O_c x_c y_c z_c$ es el sistema de coordenadas de la cámara cuyo origen es el centro óptico de la cámara. El eje z_c es el eje óptico de la cámara, los ejes $x_c y_c$ son paralelos a la dirección xy del sensor de la cámara.
- b) $O_l x_l y_l$ es el sistema de coordenadas láser cuyo origen es el centro del escáner láser. y_l es la bisectriz del ángulo del rango de exploración, y x_l es perpendicular a y_l en el plano de exploración.
- c) $O_v x_v y_v z_v$ es el sistema de coordenadas del vehículo cuyo origen es la proyección vertical del centro de escáner láser en el suelo. x_v es paralelo a x_l .

3.7.3 Calibración del sistema

El siguiente paso es la calibración del sistema. El escáner láser y la cámara tienen su propio sistema de coordenadas. Por lo tanto los puntos del láser y los datos de la cámara deben integrarse en el mismo sistema de coordenadas. En este caso, los puntos del láser se proyectan sobre el sistema de coordenadas de la imagen para su posterior procesamiento. La calibración incluye parámetros extrínsecos del escáner láser (inclinación α y altura H) y parámetros intrínsecos y extrínsecos de la cámara. La calibración intrínseca se centra en los parámetros internos de la cámara incluyendo la longitud focal, distorsión, calibración, etc. La calibración extrínseca se centra en los parámetros relativos de rotación y traslación en el sistema de coordenadas del

vehículo. Los parámetros intrínsecos son bastante estables para que la calibración intrínseca se pueda realizar con antelación. Por otro lado, el objetivo de la calibración extrínseca es encontrar los parámetros relativos de rotación y traslación entre diferentes sistemas de coordenadas.

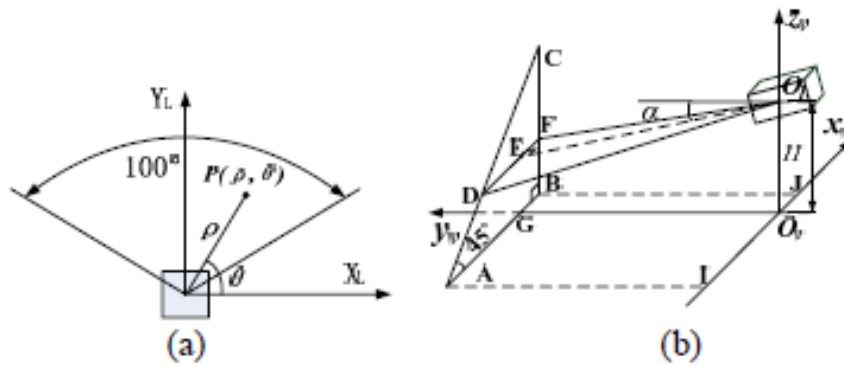


Ilustración 70. Modelo geométrico del escáner láser para la calibración extrínseca.

En la ilustración 70 se puede observar el modelo geométrico del escáner láser para la calibración extrínseca. (a) corresponde a la expresión de las coordenadas del punto P escaneados por el escáner láser. θ es el ángulo de lectura. ρ es la exploración distancia. (b) corresponde al modelo geométrico. CBA es un triángulo isósceles de calibración con un ángulo recto. Y O_lDF es el plano de exploración del escáner láser.

En primer lugar, se usa el triángulo isósceles para calibrar los parámetros extrínsecos del escáner láser. El modelo puede ser construido de la siguiente manera según la ecuación 18:

$$\begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix} = R_l \begin{bmatrix} x_l \\ y_l \\ 0 \end{bmatrix} + T_l$$

Ecuación 18

En este modelo, R_l es una matriz 3×3 de rotación. T_l es el vector de traslación. Debido a que el eje x_v es paralelo al eje x_l , el desvío es cero. La ecuación anterior se puede convertir en:

$$\begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x_l \\ y_l \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ H \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix} = \begin{bmatrix} \rho \cos \theta \\ \rho \sin \theta \cos \alpha \\ H - \rho \sin \theta \sin \alpha \end{bmatrix}$$

Ecuación 19

Ecuación 20

Donde θ es el ángulo de lectura, ρ es la distancia. Por otro lado, $DF = CF$.

$Z_{vF} = BC - DF$. En $\Delta O_l DF$, DF se puede expresar como:

$$DF = \sqrt{\rho_F^2 + \rho_D^2 - 2\rho_F\rho_D \cos(\theta_D - \theta_F)}$$

Ecuación 19

Se pueden configurar dos ecuaciones de restricción con dos puntos de restricción, H y α , se puede obtener:

$$\begin{cases} z_{wF1} = H - \rho_{F1} \sin \theta_{F1} \sin \alpha \\ z_{wF2} = H - \rho_{F1} \sin \theta_{F2} \sin \alpha \end{cases}$$

Ecuación 20

Entonces, las coordenadas polares del escáner láser pueden ser trasladadas a las coordenadas del vehículo. Los puntos del láser en el sistema de coordenadas del vehículo (x_v, y_v, z_v) pueden ser trasladados en el sistema de coordenadas de la cámara (x_c, y_c, z_c) :

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R_c \begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix} + T_c$$

Ecuación 21

Donde R_c es una matriz 3×3 de rotación, T_c es el vector de translación. De esta forma, las coordenadas del vehículo pueden ser proyectadas en el sistema de coordenadas de la imagen:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f/dx & 0 & u_0 \\ 0 & f/dy & v_0 \\ 0 & 0 & 1 \end{bmatrix} \left[R_c \begin{bmatrix} x_v \\ y_v \\ z_v \end{bmatrix} + T_c \right]$$

Ecuación 22

Donde (u_0, v_0) es el centro del píxel de la imagen del punto (u, v) , f es la distancia focal de la cámara, dx es la distancia del sensor en la dirección x , dy es la distancia del sensor en la dirección y .

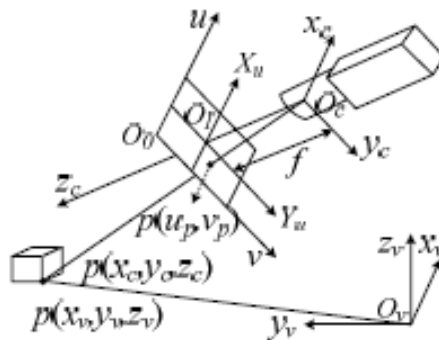


Ilustración 71. Punto de extracción de características, la relación de los tres sistemas de coordenadas.

El mismo punto P en un sistema de coordenadas diferente se puede expresar con un valor de coordenadas diferentes.

3.7.4 Fusión de datos

Una vez calibrados los parámetros del escáner láser y la cámara, el siguiente paso es integrar los puntos de láser en la imagen. Asumiendo un punto en el sistema de

coordenadas del láser P $(x_l, y_l, 0)$ se puede trasladar a (x_v, y_v, z_v) en las coordenadas del vehículo, luego se convierte en (x_c, y_c, z_c) en el sistema de coordenadas de la cámara, y, finalmente, se proyecta sobre $(u, v, 1)$ en sistema de coordenadas de la imagen:

$$(u, v, 1)^T = K_c (R_c (x_v, y_v, z_v)^T + T_c)$$

Ecuación 23

Donde K_c es una matriz 3×3 de parámetros intrínsecos. De esta forma todos los puntos del láser se han proyectado en las coordenadas de la imagen. Uno de los objetivos de la fusión de datos es el de integrar los puntos de láser e imágenes en el sistema de coordenadas del vehículo. El otro objetivo es proyectar los puntos de láser en las imágenes. Esta función es importante para la composición de la región de interés (ROI) en la imagen y además para mejorar el rendimiento de detección en tiempo real.

3.7.5 Resultados experimentales

La plataforma de prueba se ilustra a continuación:



Ilustración 72. La plataforma del sistema.

El primer paso consiste en calibrar los parámetros extrínsecos del escáner láser,

incluyendo la altura de montaje. Se coloca la placa de calibración en dos posiciones diferentes para crear dos ecuaciones de restricción (Ecuación 22). El segundo paso es extraer los puntos de la función en el láser escáner y la cámara para calibrar los parámetros de la cámara como se puede ver en la ilustración 73. El tercer paso es comprobar los resultados de calibración. Por último, el programa puede proyectar los puntos láser en la imagen como se ve en la ilustración 74.

En la ilustración 73, se han diseñado las distintas posiciones y posturas de la caja de calibración. Los puntos rojos en la imagen y en la vista de pájaro del escáner láser son los puntos de fusión.

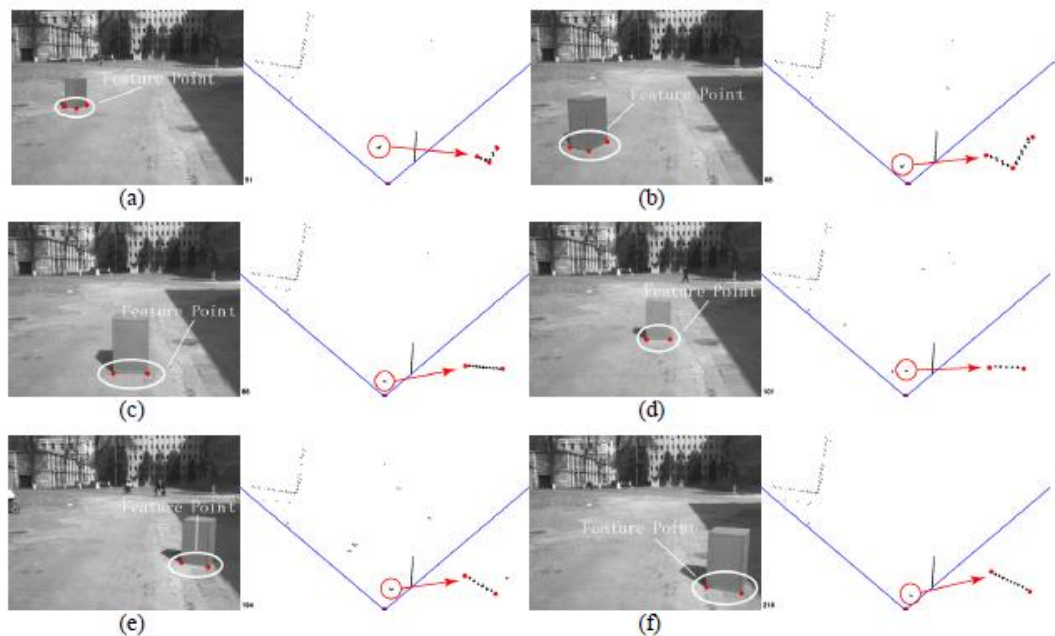


Ilustración 73. Puntos de extracción de características y proceso de calibración.

Los puntos de la imagen se corresponden con los del láser de la imagen. Y los puntos en los círculos rojos de la imagen de láser se multiplican por los puntos cerca de la flecha. Para calcular con precisión, la caja de calibración se colocó en una posición diferente en el campo de visión del conjunto de la imagen. Pero en la práctica, si la caja de calibración está colocada demasiado lejos, es difícil extraer los puntos de función en el imagen de la cámara, si está demasiado cerca, los puntos de función están fuera del campo de visión de la cámara. De acuerdo con las posiciones y

orientaciones de la caja de calibración, los resultados de calibración de los parámetros de la cámara intrínsecos y extrínsecos se muestran en la tabla 4.

Tabla 4. Resultados con los parámetros de la cámara

| | | | |
|-------------------|-------------|-------------------|-------------|
| $f(\text{mm})$ | 8.589266 | $kappa_1$ | -0.000711 |
| $T_x(\text{mm})$ | 38.034058 | $T_y(\text{mm})$ | 1833.032104 |
| $T_z(\text{mm})$ | 2041.851074 | $R_x(\text{rad})$ | 1.647960 |
| $R_y(\text{rad})$ | 0.014427 | $R_z(\text{rad})$ | -0.014106 |
| C_x | 400.093994 | C_y | 296.292130 |

Como se muestra en la ilustración 74, se ha probado el método de fusión en un entorno diferente. El ancho de la caja de calibración se puede calcular, a partir de datos de la imagen 638.25mm cuando se traduce a las coordenadas del vehículo. Y el ancho de la caja real es 640 mm, siendo el error del 2,7%. Por lo que cumple los requisitos de la aplicación práctica. La ilustración 74(a) muestra el resultado de la fusión en un entorno complejo. La ilustración 74(b) muestra el resultado de la fusión entre los límites de carreteras y dos bicicletas. La ilustración 74(c) muestra el resultado de la fusión del seguimiento del vehículo de prueba a otro vehículo de prueba. A través de las pruebas experimentales, los resultados mostraron que los puntos del láser podrían ser mostrados en la imagen claramente. El sistema fue probado en un buen número de escenarios. Se puede llegar a generar confianza en resultados y un buen desempeño. Por lo tanto es un método válido.

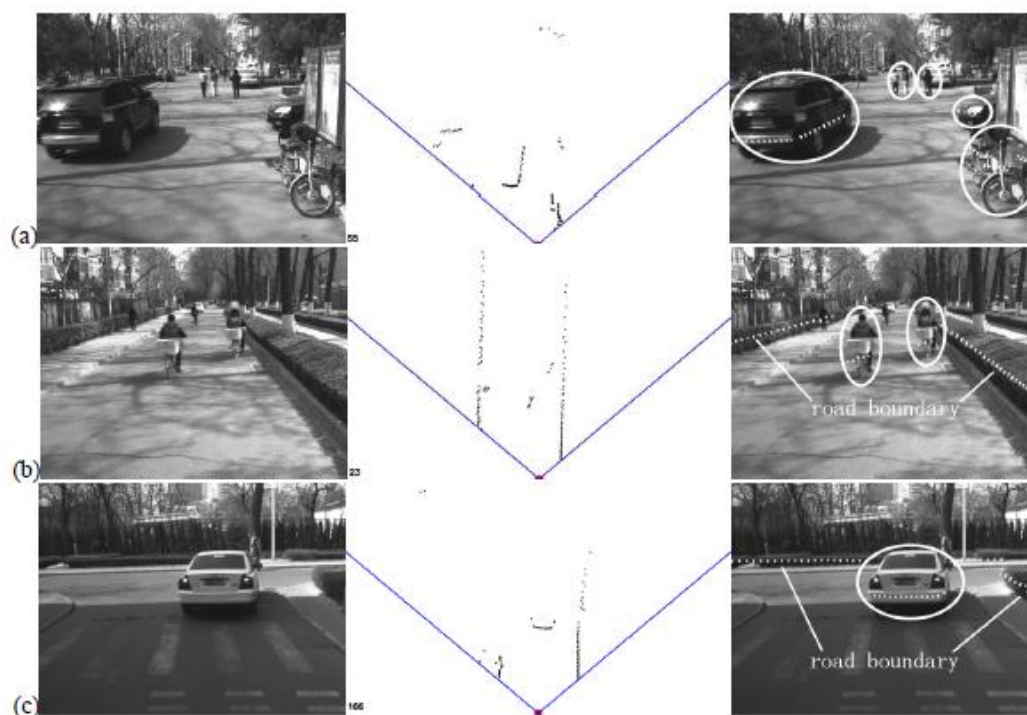


Ilustración 74. Resultados de la fusión de datos

3.8 Comparison between GMM and KDE Data Fusion methods for Particle Filtering: Application to Pedestrian Detection from Laser and Video Measurements [14]

Autores: *S. Gidel, C. Blanc, T. Chateau, P. Checchin y L. Trassoudaine. Clermont Université, Université Blaise Pascal, France.*

3.8.1 Introducción

El sistema de detección de peatones con láser tiene una fuerte capacidad de contar y realizar el seguimiento de los peatones, incluso en una situación de muy alta densidad. Sin embargo, las evidentes limitaciones de este sensor (no hay información sobre la forma, contorno, textura, el color de los objetos), su sensibilidad a las condiciones atmosféricas como la lluvia y la niebla y las oclusiones frecuentes entre objetos, requieren idear un método de fusión de láser / cámara para mejorar un sistema anticolidión de los peatones. La principal dificultad de la fusión de datos se encuentra en la asociación de las nuevas observaciones procedentes de diferentes sensores. Por lo tanto, dos problemas distintos tienen que ser resueltos en forma conjunta: la asociación de datos y la estimación. Para ello se describen dos metodologías para la realización de un BDDF (Bayesian Decentralized Data Fusion) para mantener el aspecto estocástico del conjunto de partículas. La primera es un método de transformación de la representación de partículas a una GMM (modelo de mezcla de gaussianas). El segundo algoritmo se aproxima a las partículas por una representación de KDE (estimador de la densidad Parzen).

3.8.2 Información general

El propósito de este método es realizar un seguimiento a los peatones desde un vehículo en movimiento. Para desarrollar este enfoque, la idea de emplear un filtro de

partículas surge del hecho de que los peatones no siguen un movimiento lineal ni gaussiano, por lo que el filtro de partículas parece, a priori la mejor solución.

El seguimiento con un filtro de partículas generalmente utiliza un paso de asociación de datos, en el que cada objetivo se asigna a un sensor de observación. Puesto que el conjunto de partículas no es gaussiano, se proponen dos métodos que permiten llevar a cabo una fusión de datos de dos nubes de partículas diferentes. El primer método es una transformación de la representación de partículas a la GMM. El segundo algoritmo aproxima las partículas por una representación de Parzen. En la ilustración 75 se muestra la arquitectura usada mediante la fusión de los datos del láser y la cámara de visión.

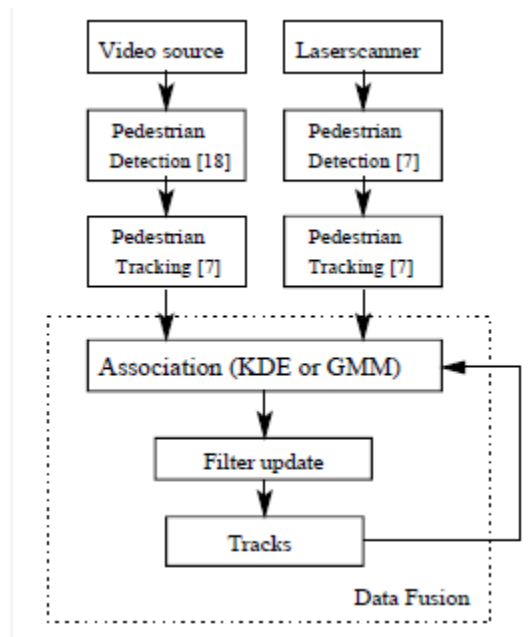


Ilustración 75. Arquitectura de la fusión de datos usando la información proporcionada por el lidar y la cámara de visión.

En cuanto a los sensores utilizados, el vehículo de prueba fue equipado con un láser XT Ibeo ALASCA montado en el centro del frontal del vehículo de prueba y una cámara de vídeo OneMarlin 146-C ubicada en la parte inferior del sensor láser.



Ilustración 76. A la izquierda, el escáner láser Ibeo ALASCA XT y la cámara Marlin 146-C. A la derecha, el vehículo de pruebas IGN.

3.8.3 Algoritmos de fusión de partículas

En el caso del filtro de partículas, la asociación de datos es un problema crucial para un mejor muestreo de partículas previa distribución. De hecho, una de las principales debilidades de los filtros de partículas es la incapacidad de explorar adecuadamente el espacio de estados si el apoyo de la distribución a priori tiene poca coincidencia con la función de verosimilitud. Una solución es colocar modelos de mezcla en las muestras, que puede tomar en cuenta de manera más eficiente la función no especificada de la distribución de las partículas predichas por el filtro de SIR. Se utilizarán modelos de mezcla y los métodos del núcleo en este con el fin de mejorar la Fusión Bayesiana de datos descentralizada (BDDF).

Por lo tanto, se va a presentar la matriz de asociación A_k para describir la asociación entre el conjunto de partículas (en representación de mediciones) Z_k y el conjunto de partículas (que representa a los objetivos) X_k . Se tiene que calcular A_k ya sea con un método paramétrico como el método de GMM o un método no paramétrico como KDE. Por último, también se define una variable aleatoria de Bernoulli $w_h \in \{w_1, w_2\}$ dada por $w_h = w_1$ si el evento asociado contiene datos fusionados, o $w_h = w_2$ en los demás casos.

3.8.3.1 Método GMM

Con el fin de una mejor aproximación a la distribución establecida de partículas, lo que permite una fusión inteligente de datos, se propone aproximar la distribución estimada con un modelo de mezcla de gaussianas (GMM). En un modelo de GMM, la distribución de probabilidad de un conjunto de partículas \mathbf{x} es una mezcla de funciones de densidad de probabilidad gaussiana M (GPDF), que se define de la siguiente manera:

$$p(\mathbf{x}|\Theta) = \sum_{m=1}^M \alpha_m p(\mathbf{x}|\theta_m)$$

Ecuación 24

donde $\theta_1, \dots, \theta_m$ son los parámetros de las distribuciones de Gauss y $\alpha = [\alpha_1, \dots, \alpha_M]$ es el vector de ponderación, de tal manera que $\sum_{m=1}^M \alpha_m = 1$. El conjunto completo de parámetros que especifican el modelo de mezcla es $\theta = (\alpha; \theta_1, \dots, \theta_M)$ con cada parámetro $\theta_m = (\mu_m, \Sigma_m)$ que consiste en un vector de medias μ y una matriz de covarianza Σ . Teniendo en cuenta que un vector de dos dimensiones $\mathbf{X} = \{\mathbf{x}_k^{1,l}, \dots, \mathbf{x}_k^{N_x,l}\}_{l=1}^N$ representa el vector compuesto por conjuntos de partículas diversas y w_h las probabilidades de cada clase que se describen como una combinación lineal de Gauss mezcla de funciones de densidad de probabilidad:

$$p(\mathbf{X}|w_h, \Theta^{i,j}) = \sum_{m=1}^M \alpha_m^{i,j} p(\mathbf{X}|\theta_m^{i,j})$$

Ecuación 25

Donde cada componente está dada por GPDF

$$p(\mathbf{X}|\Theta_m^{i,j}) = \frac{1}{\sqrt{(2\pi)^{|\Sigma_m^{i,j}|}} \exp \left[-\frac{1}{2} (\mathbf{X} - \mu_m^{i,j})^t (\Sigma_m^{i,j})^{-1} (\mathbf{X} - \mu_m^{i,j}) \right]}$$

Ecuación 26

El parámetro GMM para cada clase de objeto se calcula utilizando el algoritmo de expectativa de maximización (EM). El algoritmo EM calcula los parámetros de máxima verosimilitud en modelos estadísticos con variables que no son observadas, teniendo en cuenta momentos iniciales. El valor de las probabilidades de mezcla $x_k \in w_1$ es elegido por el estimador de máxima probabilidad como una fusión de datos:

$$A_k^{ij} | h = \arg \max_h (\alpha_m^{ij})$$

Ecuación 27

3.8.3.2 Método KDE

Para las estimaciones de densidad Parzen, cualquier tipo de núcleo se puede utilizar para representar una distribución de probabilidad. Sin embargo, los preferidos son los núcleos de Gauss, ya que la mayoría de sus operaciones son en forma cerrada y por lo tanto, eficientes. El estimador de la densidad Parzen es similar a un GMM, excepto que cada componente tiene las mismas covarianzas. La ecuación de una estimación de la densidad Parzen con un núcleo de Gauss es la siguiente:

$$p(\mathbf{x}) = \sum_{m=1}^M \gamma_m \phi(\mathbf{x}; \mu_m, \Sigma_m)$$

Ecuación 28

Donde $\phi(\mathbf{x})$ es la densidad de probabilidad gaussiana en \mathbf{x} y γ_m son los pesos donde $\sum_{m=1}^M \gamma_m = 1$. Por lo tanto se propone un enfoque de construir un modelo no paramétrico sobre la base de las funciones del núcleo, lo que permite una selección inteligente de los datos de fusión más relevantes de una función de análisis de probabilidad. Una función de probabilidad permite la clasificación de cada una de partículas como el centro de gravedad de asociación o no.

Sea $X = \{x_k\}_{k=1, \dots, N_s}$ el vector compuesto por varios conjuntos de partículas, la función de probabilidad $p(X | w_h)$ permite calcular la probabilidad de que una partícula pertenezca a una base de datos fusionados. Se propone para modelar la

probabilidad $p(\mathbf{X} | w_k)$ por un modelo no-paramétrico mediante una estimación basada en las funciones del núcleo (modelo de ventana Parzen).

$$p(\mathbf{X}|w_k) = \frac{1}{2 \cdot N} \sum_{l=1}^N \sum_{h=1}^N \varphi(\mathbf{x}_k^{i,l}, \mathbf{x}_k^{j,h})$$

Ecuación 29

Finalmente $\varphi(x_k^{i,l}, x_k^{j,h})$ es la función del núcleo que permite modificar la zona de influencia de un punto con sus vecinos, se define por:

$$\varphi(\mathbf{x}_k^{i,l}, \mathbf{x}_k^{j,h}) = \exp[-\lambda_c \cdot d_c(\mathbf{x}_k^{i,l}, \mathbf{x}_k^{j,h})]$$

Ecuación 30

El parámetro λ_c permite ajustar los pesos. La distancia d_c utiliza una distancia de *Mahalanobis* definida por:

$$d_c(\mathbf{x}_k^{i,l}, \mathbf{x}_k^{j,h}) = (\mathbf{x}_k^{i,l} - \mathbf{x}_k^{j,h}) \Sigma_\varphi^{i,j-1} (\mathbf{x}_k^{i,l} - \mathbf{x}_k^{j,h})^T$$

Ecuación 31

con Σ_φ^i y Σ_φ^j . La matriz de covarianza dada por los algoritmos de seguimiento representa las incertidumbres de la posición de los peatones.

$$\Sigma_\varphi^{i,j} = \Sigma_\varphi^i + \Sigma_\varphi^j$$

Ecuación 32

La partícula en 2D con la más alta probabilidad $x_k \in w_1$ es elegida por el estimador de máxima verosimilitud como datos de fusión:

$$A_k^{i,j} | h = \underset{h}{\operatorname{argmax}} (p(\mathbf{X} | w_h \in w_1))$$

Ecuación 33

3.8.4 Resultados

Los resultados de varios de los datos de la fusión del láser y vídeo son los siguientes. Como el láser y los datos de la cámara no tienen el mismo fotograma de referencia se elige el fotograma de referencia del láser para la fusión.



Ilustración 77. Ejemplo de detección, en una sección transversal después de una fusión centralizada de datos del lidar y de la imagen. Se detectan peatones en diferentes orientaciones.

En la imagen anterior se puede observar cómo se detectan los peatones con el vehículo de pruebas moviéndose a una velocidad que va de 0 a 50 km / h, lo que permite poner a prueba la robustez de este método. El láser puede generar falsos positivos que se solucionarían con la fusión con la cámara de visión. Finalmente se confirma una mejor aproximación de los métodos de GMM en un el seguimiento de los peatones con un número menos elevado de partículas como se muestra en las siguientes figuras:

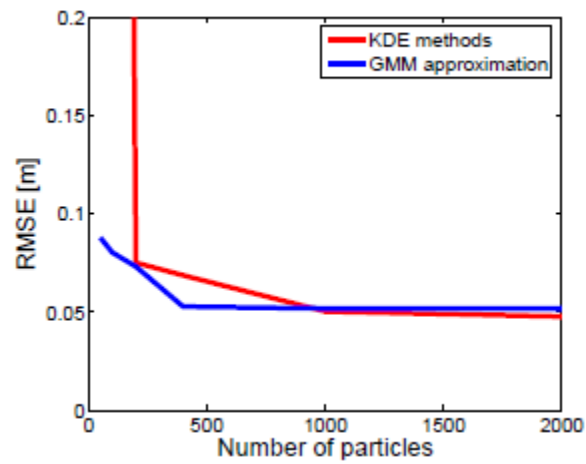


Ilustración 78. Comparativa del error medio cuadrático en posición radial entre los métodos de GMM y KDE, cuando va aumentando el número de partículas.

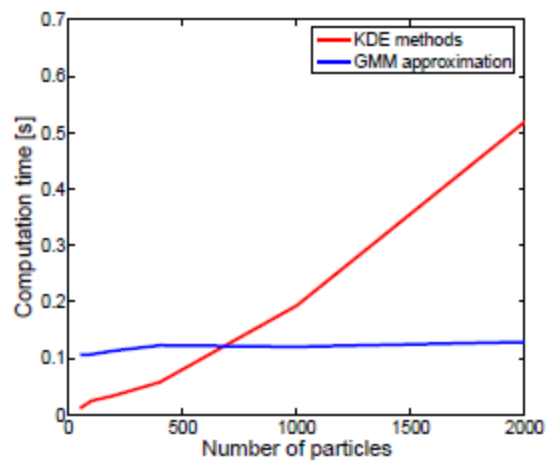


Ilustración 79. Comparación del tiempo de cálculo entre los métodos GMM y KDE (en el caso de dos sensores), cuando el número de partículas va aumentando

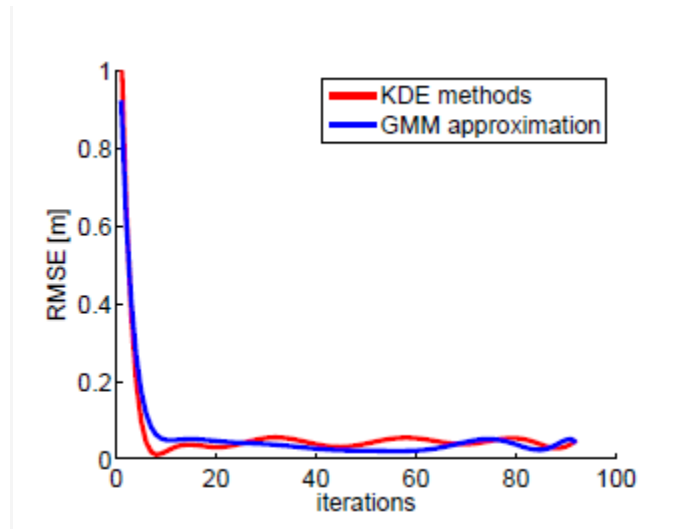


Ilustración 80. Comparativa del error medio cuadrático en posición radial entre los métodos de GMM y KDE en el seguimiento de un peatón presentado en la figura. 77.

No obstante el método KDE es más preciso que la aproximación GMM cuando aumenta el número de partículas (más de 1000 partículas). Sin embargo, cuando el número de partículas disminuye, el método de GMM se vuelve más fiable.

3.9 Fast Pedestrian Detection in Dense Environment with a Laser Scanner and a Camera [15]

Autores: G. Gate, A. Breheret y F. Nashashibi. Robotics Laboratory, Mines ParisTech, Paris.

3.9.1 Introducción

En este método se desarrolla un algoritmo que se basa en la detección, seguimiento y clasificación de posibles peatones mediante una imagen y hace uso de un algoritmo basado en la visión por computador para refinar la clasificación de los obstáculos detectados como personas. Para ello se utiliza un láser y una cámara de visión en blanco y negro.

3.9.2 Principios del algoritmo

Para la realización de este algoritmo se propone calcular tres probabilidades que se utilizarán para decidir qué objetos detectados son personas. Por cada blanco en seguimiento i en tiempo k , el sistema espera obtener estimaciones de:

- La probabilidad $P(D_k^i | Z_{0:k})$ del evento D_k^i : "el objetivo es un obstáculo real" (llamada la señal de detección).
- La probabilidad $P(R_k^i | D_k^i, Z_{0:k})$ del evento R_k^i : "el objetivo es una persona" dada D_k^i (llamada la señal de reconocimiento).
- La probabilidad $P(T_k^i | D_k^i, Z_{0:k})$ del evento T_k^i : "el objetivo está seguido correctamente", dado D_k^i (llamada la señal de seguimiento).

Para estas probabilidades se calcula cada algoritmo de los sensores láser y de visión por separado y se combinan para generar estimaciones consolidadas.

3.9.3 Escáner láser

A continuación se explica más detalladamente el algoritmo utilizado para el escáner láser:

A. Los objetos de detección

La imagen proporcionada por el rango de un escáner láser es de un número de puntos fijos que deben ser agrupados en "objetos". Lo primero es descartar aquellos objetos tales como automóviles, camiones y paredes, objetos "rectilíneos". Para simplificar utilizamos una nube de puntos y agrupamos los puntos con menos distancia entre sí en segmentos.

B. Cálculo de las oclusiones

Los contornos reales de un obstáculo por lo general sólo se ven parcialmente a través del escáner láser. Debido a la oclusión, el tamaño visible como medida después de que se haya realizado el proceso de agrupamiento puede ser engañoso.

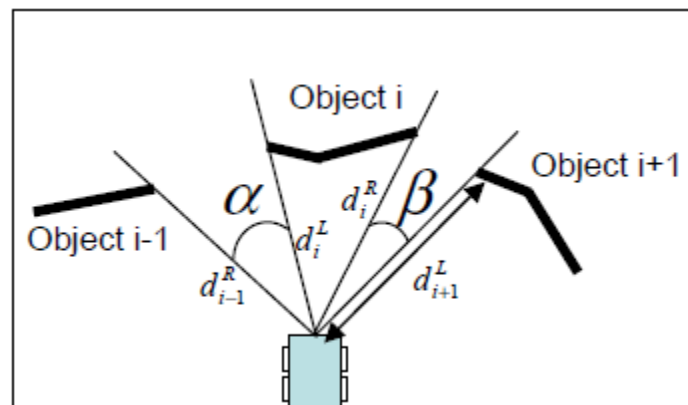


Ilustración 81. Para decidir si el objeto i se ocluye o no, se considera la posición de los dos obstáculos que le rodean (objeto $i-1$ y objeto $i+1$).

Por eso, usando la notación de la figura anterior, cada "objeto" es etiquetados como "ocluido" o "no ocluido" siguiendo la regla:

- Si $(\alpha < \text{resolución}) \wedge (d_{i-1}^R < d_i^L) \vee (\beta < \text{resolución}) \wedge (d_i^R < d_{i+1}^L)$ es cierto, entonces el objeto i se ocluye.
- De lo contrario no se ocluye.

El término "resolución" se refiere a la resolución angular del escáner láser

C. Seguimiento y la Asociación de datos

La forma en que las nuevas medidas están asociadas a los objetivos existentes es un punto clave en todo sistema de percepción. Cada nueva medida está asociada a la más cercana con la distancia de *Mahalanobis* y siempre esta distancia es menor que un parámetro α (gating). Suponiendo que la densidad de probabilidad del vector de estado aleatorio (X, Y, V_x, V_y) y los ruidos de los sistemas son gaussianos, se puede utilizar un filtro de Kalman para estimar el vector de estado a través de un modelo de velocidad constante.

D. Cálculo de las probabilidades de clasificación

Como se mencionó anteriormente, la clasificación de un objetivo como "Peatón" se hace sobre la base de los valores tomados por las tres notas de la clasificación. Para facilitar la estimación de estas notas, se utiliza un filtro Bayesiano de la siguiente manera:

$$P(D_k^i | Z_{0:k}) \propto P(z_k | D_k^i) P(D_{k-1}^i | Z_{0:k-1})$$

$$P(T_k^i | D_k^i, Z_{0:k}) \propto P(z_k | T_k^i, D_k^i) P(T_{k-1}^i | D_{k-1}^i, Z_{0:k-1})$$

$$P(R_k^i | D_k^i, Z_{0:k}) \propto P(z_k | R_k^i, D_k^i) P(R_{k-1}^i | D_{k-1}^i, Z_{0:k-1})$$

Como resultado, para cada paso de tiempo, una nota de la clasificación es la combinación de la estimación anterior y una probabilidad del cálculo que se detalla a continuación.

1) El cálculo de la señal de detección

El cálculo de la probabilidad asociada con la señal de detección se basa en el número de puntos $N_{observed}$ que se agrupan en los objetos asociados con ese objetivo. De hecho, un gran número de puntos es probable que indique un verdadero obstáculo, mientras que la medida de un grupo con sólo algunos puntos podría ser errónea. Para normalizar esta estimación se utiliza el número de puntos de N_{max} que observaría si el objeto que ha detectado el rango del sensor tiene el tamaño mínimo.

$$P(z_k | D_k) \cong \frac{N_{observed}}{N_{max}}$$

Ecuación 34

2) El cálculo de la señal de seguimiento

La señal de seguimiento se calcula sobre la base de la traza de la matriz de covarianza P_k del Filtro de Kalman utilizando la siguiente fórmula de la ecuación 37, donde P_{final} es el valor teórico de P_k después de la convergencia del Filtro de Kalman.

$$P(T_k | Z_{0:k}) \cong \sqrt{\frac{trace(P_{final})}{trace(P_k)}}$$

Ecuación 35

Para simplificar, P_{final} se calcula experimentalmente (el ruido de matrices en las ecuaciones de Kalman se establece en un valor fijo).

3) Cálculo de la señal de reconocimiento

La estimación de la probabilidad asociada con el reconocimiento se basa en tres criterios: el tamaño del objeto, si está ocluido y su velocidad, como se detalla a continuación.

$$P(z_k | R_k, D_k) \cong \Gamma(s_k^i) \times \Theta(occluded_k^i) \times \Omega(v_k^i)$$

Ecuación 36

- a) El tamaño global S_k de los grupos asociados se define como la diagonal del rectángulo:

$$\Gamma(s_k^i) = \begin{cases} 0 & \forall s_k^i < a_1, s_k^i \geq a_4 \\ \frac{s_k^i - a_1}{a_2 - a_1} & \forall a_1 \leq s_k^i < a_2 \\ 1 & \forall a_2 \leq s_k^i < a_3 \\ \frac{a_4 - s_k^i}{a_4 - a_3} & \forall a_3 \leq s_k^i < a_4 \end{cases}$$

Ecuación 37

- b) Si los grupos están ocluidos:

$$\Theta(occluded_k^i) = \begin{cases} 1 & \text{if } occluded_k^i = 0 \\ 0 & \text{if } occluded_k^i = 1 \end{cases}$$

Ecuación 38

- c) Si la velocidad del objetivo es realista para un peatón:

$$\Omega(v_k^i) = \begin{cases} 1 & \text{if } v_k^i \leq v_{\max} \\ 0 & \text{else} \end{cases}$$

Ecuación 39

3.9.4 Sistema de visión

Cada objetivo enviado por el algoritmo del escáner láser que se ha descrito anteriormente se proyecta en la imagen calibrada de la cámara. Después las regiones de interés son procesadas por el algoritmo AdaBoost.

A. Búsqueda en la región de interés

Debido al posible balanceo y cabeceo del robot, puede haber un desplazamiento entre la posición de la región prevista y real del obstáculo. Para evitar este problema, el algoritmo de reconocimiento se realiza en una zona cercana de la región proyectada de interés. Para mantener el tiempo de cálculo razonable, se procesan unas 10 imágenes para cada objetivo enviado por el algoritmo del escáner láser.

B. Visión basada en la clasificación

Como el algoritmo AdaBoost únicamente permite distinguir un peatón de un objeto no humano, no se puede calcular la detección ni el seguimiento de los mismos. Sin embargo, la señal de reconocimiento está dada directamente por la suma ponderada de los clasificadores débiles u_k^i :

$$P(R_k^i | D_k^i, Z_{0:k}) \cong u_k^i$$

Ecuación 40

3.9.5 Fusión de datos

Una vez calculados los algoritmos del escáner láser y del sistema de visión se procede a combinarlos mediante la fusión de ambos:

A. Combinación de las señales de la clasificación

Con el fin de clasificar finalmente los objetivos como ser humano o no humano, la señal de reconocimiento estimada realizada por el algoritmo del escáner láser tiene que combinarse con la estimación del algoritmo de visión. En cuanto a la señal

de reconocimiento, se utiliza el marco Bayesiano para calcular una estimación global de $P(R_k^i | D_k^i, Z_{0:k})$ a partir de:

- La estimación del algoritmo del escáner láser $P(R_k^i | D_k^i, Z_{0:k}, A_1)$
- Y la estimación del algoritmo de visión $P(R_k^i | D_k^i, Z_{0:k}, A_2)$

Los dos eventos A_j pueden ser entendidos como “el algoritmo j da una estimación válida”. Es de suponer que estos dos eventos son independientes. En la siguiente deducción, el evento “ R_k^i dado D_k^i y $Z_{0:k}$ ” se nombra para mayor claridad como C_k^i . La siguiente ecuación derivada de las reglas básicas de probabilidad se utiliza para combinar las dos estimaciones de la señal de reconocimiento.

$$\begin{aligned} P(C_k^i) &= P(C_k^i \cap (A_1 \cup A_2)) + P(C_k^i \cap \overline{(A_1 \cup A_2)}) \\ &= P(C_k^i | A_1)P(A_1) + P(C_k^i | A_2)P(A_2) \\ &\quad - P(C_k^i | A_1 \cap A_2)P(A_1)P(A_2) + P(C_k^i | \overline{A_1} \cap \overline{A_2})P(\overline{A_1})P(\overline{A_2}) \end{aligned}$$

Ecuación 41

Los términos de esta fórmula se interpretan de la siguiente manera. Los términos $P(C_k^i | A_j)$ son las dos señales de reconocimiento de la estimación de los dos subsistemas. $P(A_j)$ se refiere a la probabilidad de que el subsistema j pueda devolver una estimación errónea debido por ejemplo a la intemperie o condiciones de iluminación. $P(C_k^i | \overline{A_1} \cap \overline{A_2})$ se puede aproximar a $P(R_k^i)$, la probabilidad a priori de un obstáculo de ser una persona. Finalmente, la probabilidad de C_k^i verifica que las estimaciones son válidas y se aproxima como la suma ponderada:

$$P(C_k^i | A_1 \cap A_2) = \frac{P(C_k^i | A_1)P(A_1) + P(C_k^i | A_2)P(A_2)}{P(A_1) + P(A_2)}$$

Ecuación 42

B. Clasificación final

La clasificación final se hace sobre la base de un período de tres umbrales (uno por cada señal de clasificación). Los resultados que se presentan a continuación se obtienen con los siguientes límites y parámetros:

Estas son las tres condiciones que validan a un peatón:

- Señal de detección > 0.4
- Señal de reconocimiento de peatones > 0.5
- Señal de seguimiento > 0.5

C. Resumen del algoritmo

Esta estrategia de fusión se basa en la combinación de las estimaciones de los dos algoritmos basados en los sensores que están contribuyendo. Mientras que el algoritmo del escáner láser genera una estimación de las tres señales de clasificación, el sistema de visión sólo es capaz de proporcionar una estimación de la señal de reconocimiento. Por lo tanto se combinan mediante las anteriores ecuaciones para obtener la clasificación final.

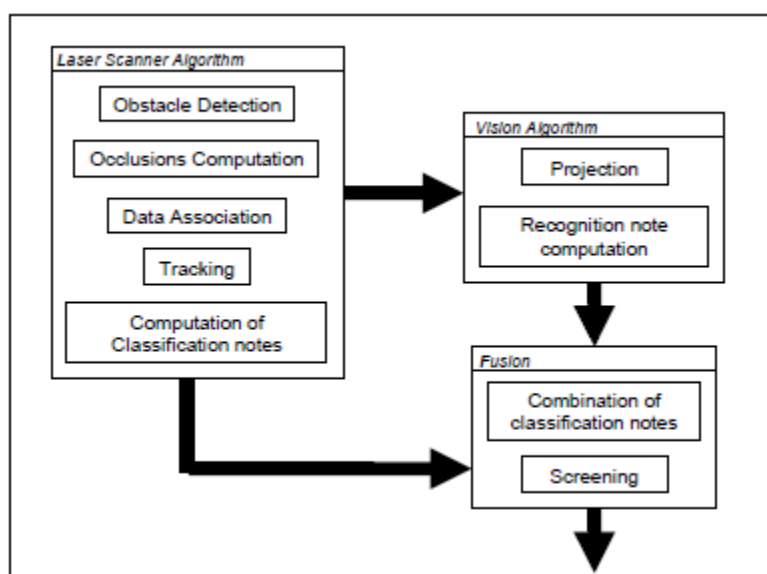


Ilustración 73. Principios del algoritmo de fusión centralizada y basada en la rápida detección de peatones en entornos densos.

3.9.6 Resultados

Como ejemplo para demostrar la validez del método empleado, en las siguientes ilustraciones se muestran los resultados obtenidos por el escáner láser, el sistema de visión y finalmente la fusión de ambos, se muestran los peatones que se detectan en un rectángulo y así se comprueba la validez de los algoritmos descritos anteriormente.

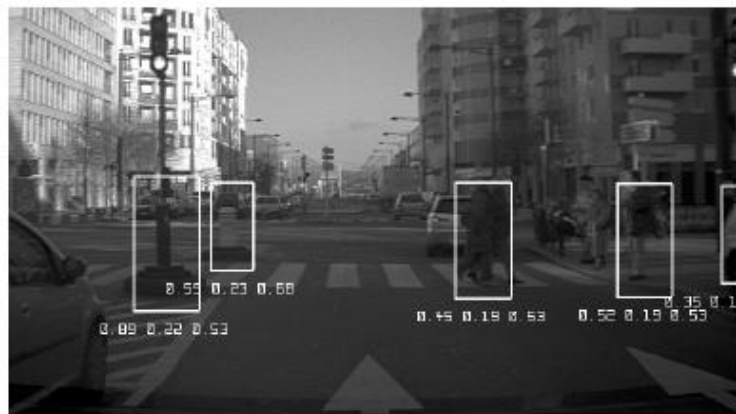


Ilustración 82 En esta figura se han calculado las señales de detección, reconocimiento y seguimiento del algoritmo del escáner láser.

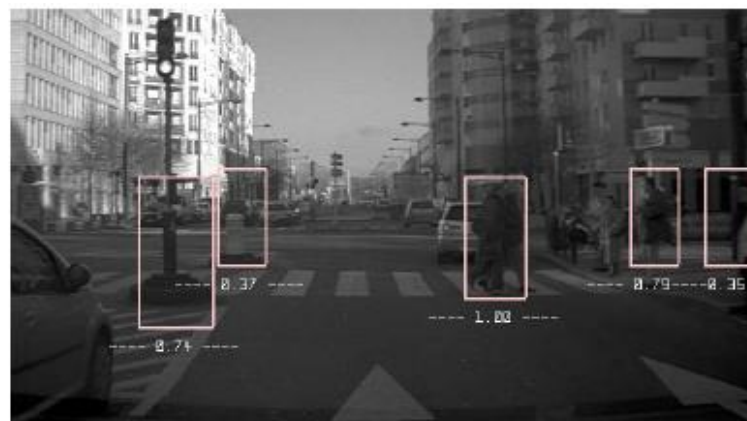


Ilustración 83 Aquí solamente se calcula la señal de seguimiento, en el algoritmo de visión.



Ilustración 84. Finalmente, mediante la fusión de ambos se pueden detectar, reconocer y seguir eficazmente los peatones de la escena.

3.10 Fusing Vision and Lidar – Synchronization, Correction and Occlusion Reasoning [16]

***Autores:** Sebastian Schneider, Michael Himmelsbach, Thorsten Luettel y Hans-Joachim Wuensche. Institute for Autonomous Systems Technology (TAS), University of the Bundeswehr Munich, Germany.*

3.10.1 Introducción

En este método se propone una fusión de bajo nivel de un sensor lidar y una cámara de visión, ya que esto nos permite utilizar la información de color y profundidad combinada en la fase de detección de objetos. Como va dirigido a escenas dinámicas hay que asegurar que los dos sensores observan el mismo espacio al mismo tiempo. Por ello hay que sincronizar los dos sensores, teniendo en cuenta el movimiento del vehículo.

3.10.2 Adquisición de datos, corrección y sincronización

El sistema de visión del vehículo de pruebas consistirá en una cámara a color que se monta sobre una plataforma giratoria cerca del espejo retrovisor, mientras que el lidar irá colocado en el techo del vehículo.

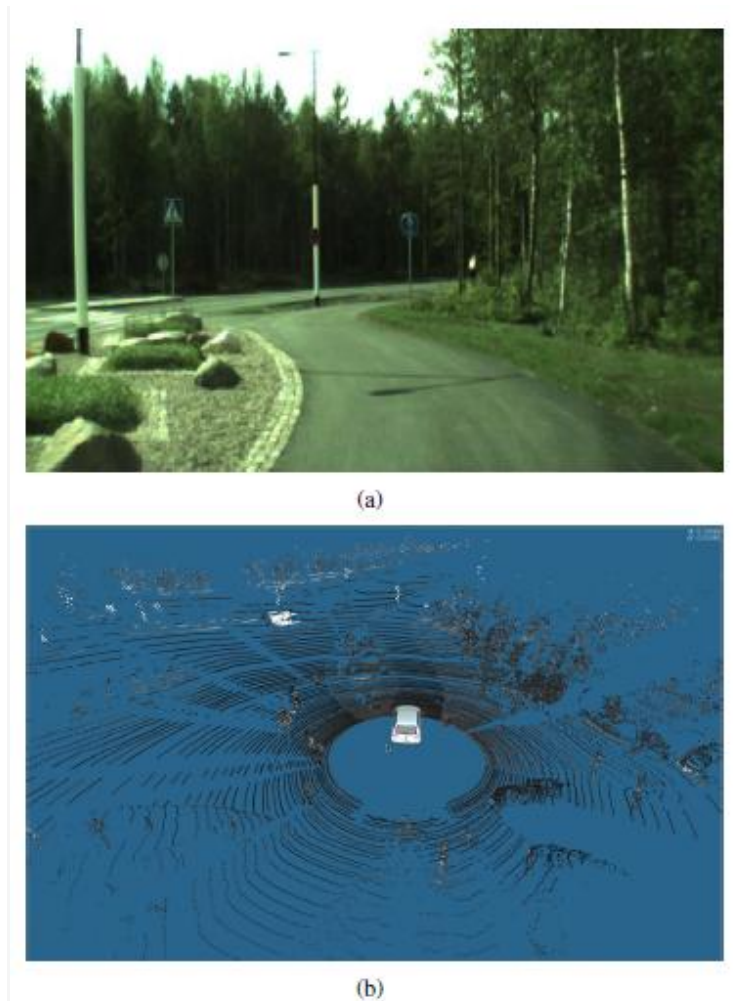


Ilustración 85. . La imagen en color de la cámara (a) y la correspondiente nube de puntos corregida de 105 puntos de datos del lidar Velodyne (b).

Antes de poder ejecutar ambos algoritmos sobre las medidas tomadas se tiene que compensar el movimiento del vehículo. Al comparar el tiempo necesario para la adquisición de datos de cada sensor, la cámara es más rápida que el escáner láser en un factor de 2. Al tener un entorno dinámico, con objetos en movimiento, si los sensores fueran asíncronos capturarían la escena en diferentes momentos, causando problemas en la fusión de datos. Por lo tanto, es de desear que tanto la cámara como el láser tomen los datos al mismo tiempo y para ello es inevitable la sincronización de los dos sensores. Mientras que la sincronización de la posición es relativamente fácil, en el tiempo es más complejo ya que la posición y el ángulo de la plataforma de la cámara, además del tiempo de exposición de la cámara, cambian con el tiempo. Por lo que un tiempo de disparo fijo podría causar errores en fusión de datos. La mejor sincronización se logra al cambiar el tiempo de exposición hasta el

momento en el que los rayos láser pasan el eje x de S_{cam} , como se muestra en la ilustración 86:

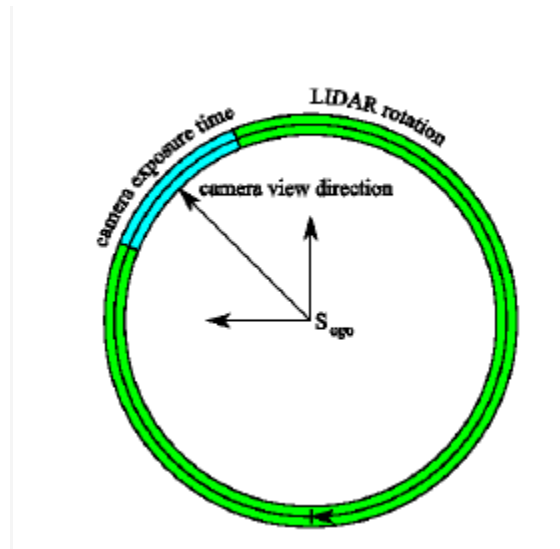


Ilustración 86. Sincronización de la cámara y el lidar. La cámara tiene que ser disparada de tal manera que capture la imagen cuando los rayos láser pasen el campo de visión de la cámara (azul). En este caso se muestra cuando el vehículo está parado.

Si el vehículo está en movimiento, la fusión de las mediciones con el láser y la cámara de visión se tiene que hacer dentro de un sistema de coordenadas diferentes, $S_{ego,1}$, describiendo la posición y orientación del vehículo en el momento de la captura de la imagen. Por lo tanto, es necesaria una transformación en los ejes de coordenadas, de $S_{ego,0}$, a $S_{ego,1}$, antes de que el algoritmo de fusión pueda funcionar correctamente. A continuación, los datos fusionados tienen que ser transformados de tal manera que las coordenadas hagan referencia con respecto a otro eje $S_{ego,2}$, al final de una revolución. Esto se puede ver en la ilustración 87:

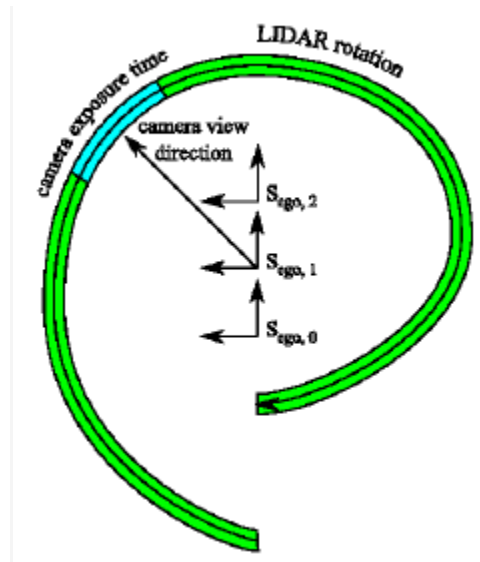


Ilustración 87. Sincronización de la cámara y el lidar. La cámara tiene que ser disparada de tal manera que capture la imagen cuando los rayos láser pasen el campo de visión de la cámara (azul). En este caso se muestra cuando el vehículo está en movimiento.

3.10.3 Fusión de datos

Después de haber sincronizado los dos sensores el siguiente paso es la fusión real de los datos. El resultado de esta fusión son datos en 3D e información de la distancia, la reflectancia así como información de color correspondiente con dichos puntos del lidar.

3.10.3.1 Transformación y asignación de color

Con el fin de proporcionar los datos de un sensor a otro sensor, tanto las posiciones como las orientaciones de los sensores se conocen. Se sabe por la calibración, la posición de los sensores con respecto al centro de gravedad del vehículo S_{ego} .

Al comparar el campo de visión de la cámara con el escáner lidar, vemos que sólo una pequeña parte de los puntos 3D se pueden ver en la imagen de la cámara.

Debido a la ubicación física en el vehículo donde se montan los sensores, sólo pueden ser visibles los puntos con la coordenada x mayor o igual a 0.

Se utiliza la matriz de transformación 4x4 A_{ego}^{img} para proyectar los puntos del plano de la imagen que se encuentra en el centro de:

$$p_j^{img} = A_{ego}^{img} \cdot p_j^{ego},$$

Ecuación 43

El superíndice indica el nombre del sistema de coordenadas a las que se hace referencia a un punto. El índice j se repite en todos los elementos de P . Hay que recordar que A_{ego}^{img} está sujeta a cambios cada vez que se mueve la plataforma de la cámara. Así que A_{ego}^{img} se tiene que actualizar con cada imagen de la cámara.

Una vez que se han calculado las coordenadas en píxeles hay que almacenar las coordenadas correspondientes en 3D en una estructura de la imagen.

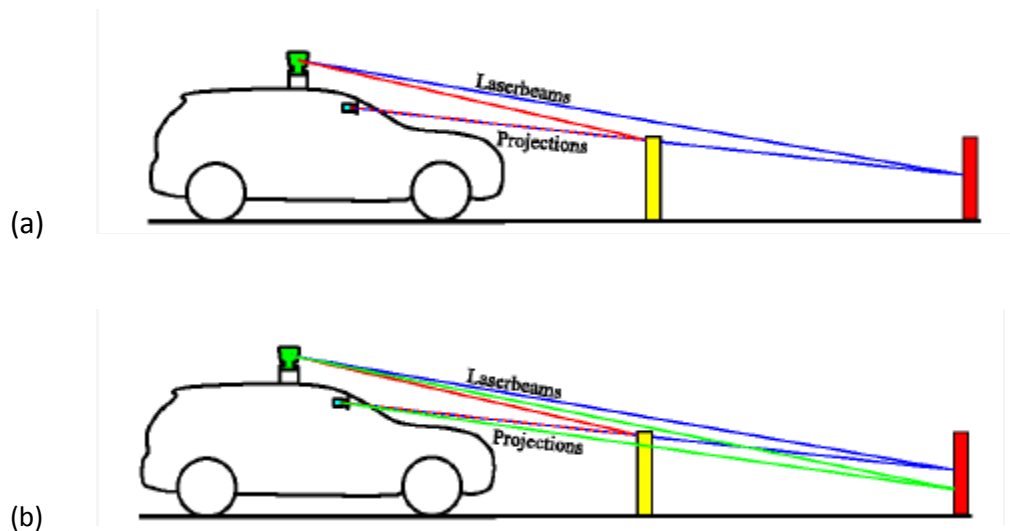


Ilustración 88. Algunos rayos láser alcanzan objetos en posiciones que se ocuyen a la cámara

En la ilustración 88 se puede observar que el láser está colocado en una posición más elevada que la cámara, por lo tanto, podrá detectar objetos (en este caso de color rojo) que para la cámara están ocultos por otro objeto (el de color amarillo). Así, aunque desde la cámara el objeto rojo esté oculto puede ser detectado mediante el láser (rayo verde).

3.10.3.2 Tratamiento de la oclusión

La fusión de los datos es correcta si los dos sensores observan el entorno desde el mismo punto de vista. Por desgracia, en el vehículo el lidar está un metro por encima de la cámara. Esto es suficiente para que el lidar pueda percibir el entorno detrás de los objetos que obstruyen la vista de la cámara. Este problema se conoce como problema de oclusión. Como el punto en el objeto amarillo está más cerca de S_{cam} , en comparación con el punto en el objeto de color rojo, el punto amarillo ocluye el punto rojo. En tal caso, sólo el punto en el objeto amarillo debe estar coloreado. Si añadimos otro rayo láser como en (b), el punto en el objeto rojo que se origina en el rayo láser verde no debe estar coloreado, ya que se obstruye. Este problema es más difícil de resolver, ya que los puntos ocluidos se proyectan de forma adyacente los puntos sin oclusión. Para solucionar este problema, se utiliza el algoritmo de segmentación de nube de puntos. Dada una nube de puntos en 3D todavía sin color, esta segmentación separa el plano de tierra y luego realiza una agrupación en 3D de los puntos restantes asignando a cada punto 3D de la misma categoría un identificador de segmento común. Esta asociación de cada punto en 3D para su segmento se mantiene intacta después de la proyección de la imagen de la cámara. El conjunto P^{img} es el conjunto de todas las proyecciones de puntos 3D que pasa la comprobación de la visibilidad. Entonces, la segmentación permite una partición del P^{img} en el subconjunto de todos los puntos cuyo origen se encuentra en plano de tierra P_{ground}^{img} y los subconjuntos P_i^{img} de los puntos que pertenecen al mismo segmento en 3D:

$$P^{img} = P_{ground}^{img} \cup \bigcup_i P_i^{img},$$

Ecuación 44

donde i se repite en todos los segmentos.

La comprobación de la oclusión consiste en los siguientes pasos (se pueden ver en la siguiente ilustración):

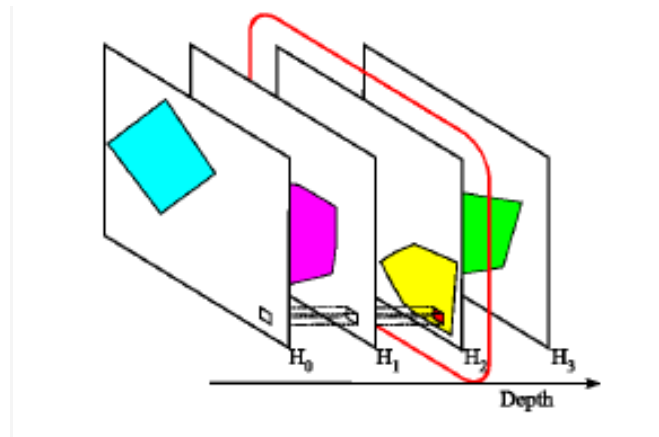


Ilustración 89. Detección de la oclusión: por cada punto que se itera a través de la profundidad de lista ordenada de los cascos convexos H_3 . Identificación del segmento del primer casco que incluye un punto (en este caso H_2) se compara con el punto de identificación del segmento.

Para cada P_i^{img} hay que iterar a través de elementos ordenados de H hasta que se encuentre con la primera $H_j \in H$, donde P_i^{img} se encuentra en el interior de la envolvente convexa. P_i^{img} es visible si el segmento id de P_i^{img} y H_j combinan, de lo contrario P_i^{img} se ocluye. Si P_i^{img} no se encuentra dentro de H_j entonces P_i^{img} forma parte del plano de tierra no ocluido y por lo tanto, visible.

3.10.4 Evaluación y resultados

Para la evaluación del algoritmo de fusión de datos y detección de oclusión se ha grabado una secuencia del vehículo por una carretera con césped y árboles a ambos lados y un cartel amarillo (0,4 m x 0,3 m) en la hierba alta en la derecha. La secuencia se inicia con la señal a una distancia de aproximadamente 26 metros y termina a una distancia de aprox. 6m del vehículo. En la ilustración 90 se muestra una imagen de muestra de la secuencia grabada. Para cada uno de los 25 *frames* han sido etiquetados los puntos 3D que pertenecen al cartel amarillo, lo que ha permitido contar el número de puntos que han sido asignados erróneamente a la señal de color amarillo.



Ilustración 90. Escena observada por ambos sensores.

La ilustración 91 muestra un diagrama de los resultados de dos test de conducción. En la primera (línea verde) la detección de la oclusión estaba inhabilitada. En el segundo (trama de color rojo) la detección de la oclusión se ha habilitado. El magenta y el azul representan las curvas de un ajuste de mínimos cuadrados (LSF) a través de los datos adquiridos.

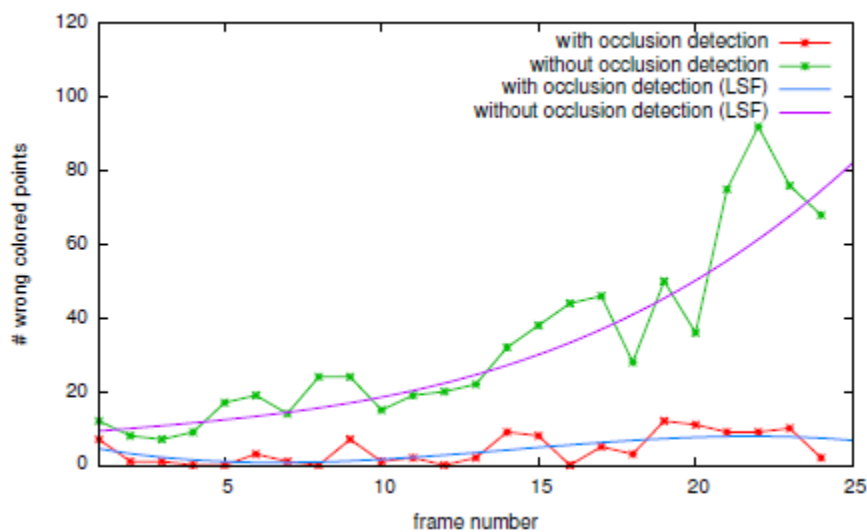


Ilustración 91. El número de píxeles de la señal mal coloreados con y sin detección de la oclusión

Los resultados, sin el detector de oclusión son los esperados. El número de puntos en 3D erróneos incrementa a la distancia en la que la señal se hace más pequeña debido a que cada vez más rayos láser adquieren las mediciones de la hierba alta que hay detrás de la señal. Comparado con eso, con el detector de oclusión

activado, la cantidad de píxeles de color erróneos es relativamente pequeña a lo largo de la secuencia.

Para finalizar, este sistema de fusión de sensores ofrece información en profundidad de imágenes de la cámara alcance la precisión muy cercana al lidar así como la reflexión de la información, por lo tanto es un método eficiente para detectar objetos ocluidos.

Capítulo 4.Descripción del sistema a desarrollar

4.1 Descripción general

Una vez realizado el estudio de los diferentes métodos existentes para realizar la detección y el seguimiento de peatones en un entorno de conducción mediante técnicas de fusión sensorial, se va a proceder al diseño de un método propio para su posterior implementación y realización de pruebas reales.

En primer lugar, necesitamos diseñar un sistema ADAS (Advance Driver Assistant Systems) que permita la correcta detección de un peatón desde un vehículo en movimiento y una vez detectado proceder a realizar un seguimiento del mismo y tratar de averiguar cómo se va a mover (predicción de la trayectoria), y una vez hecho esto utilizar un algoritmo de alerta que avise al conductor en el caso de que el peatón esté demasiado cerca. A su vez el sistema tiene que ser capaz de detectar peatones que están parcialmente ocluidos y realizar una frenada automática si fuera necesaria.



Ilustración 92. En esta imagen se puede ver cómo un peatón ocluido por un autobús intenta cruzar la calle.

Toda la programación que esto conlleva se realizara en el lenguaje de programación de Matlab, junto con librerías de libre distribución como la toolbox

Rebel, para realizar las estimaciones mediante el filtro de Kalman y otros filtros de estimación.

Para ello necesitamos utilizar la técnica de la fusión de datos de dos sensores y un posterior algoritmo para tratar dichos datos. Los sensores que se utilizarán serán los que disponemos en el laboratorio de la Universidad Carlos III, un escáner láser y una cámara de visión estéreo, que irán acoplados al vehículo de pruebas.

4.2 Descripción del Hardware utilizado

4.2.1 Vehículo de pruebas

Uno de los vehículos disponibles en el laboratorio de Ingeniería de Sistemas y Automática de la Universidad Carlos III, es el IVVI 2.0 (Vehículo inteligente basado en información visual), un Nissan Note equipado con ordenadores de a bordo y diversos sensores, que lo convierten en un coche “inteligente”, necesarios para el desarrollo de diferentes sistemas ADAS.



Ilustración 93. Vista lateral del vehículo de pruebas IVVI 2.0.

4.2.2 Sensores

A continuación se va a proceder a detallar cada uno de los elementos del vehículo de pruebas que se van a utilizar.

4.2.2.1 Sensor láser

El láser disponible en el laboratorio que se va a utilizar es un LMS SICK 291 [17] [18] (ilustración 94). Este sensor es la última variante de los telémetros láser de SICK, el cual viene en un color beige característico. El LMS SICK 291 funciona a una frecuencia de 75 Hz con un rango de exploración de 180° , con $0,25^\circ$ de resolución angular. Su rango de detección es de 30 metros (10% de reflectividad, alcance máximo de 80 metros), con un error de unos 10 mm. Debido a su robusta caja metálica, el LMS 291 es un poco voluminoso (18,5 x 15,6 x 21,0 cm), con un peso de aproximadamente 4,5 kg. Se requieren aproximadamente 20 vatios de potencia, que operan mediante un suministro de 24 VCC.



Ilustración 94. Telémetro láser SICK LMS 291

El SICK LMS 291 se basa en un espejo giratorio y un diodo láser para realizar el cálculo del tiempo de vuelo. En un desmontaje del telémetro láser, se puede ver

claramente los componentes principales, tales como el mecanismo de espejo giratorio, diodos láser, etc.



Ilustración 95. Vista interna del SICK LMS 291.

Para obtener los datos con los que se va a trabajar se va a emplear una configuración a $0,25^\circ$ de resolución y una distancia de 80 metros, con un ángulo de visión de 100° . Para lograr la resolución deseada el láser realiza cuatro escaneos independientes de un grado de resolución cada uno, cada uno de ellos separado $0,25^\circ$ con respecto al anterior escaneo. Así, es posible conseguir una resolución de $0,25^\circ$ una vez que se han realizado los cuatro escaneos, como se puede observar en la ilustración 88.

Este comportamiento permite identificar los objetos en movimiento debido al patrón que deja a lo largo de las cuatro rotaciones consecutivas. Este patrón será proporcional al movimiento del objeto y su trayectoria a lo largo de las cuatro rotaciones.

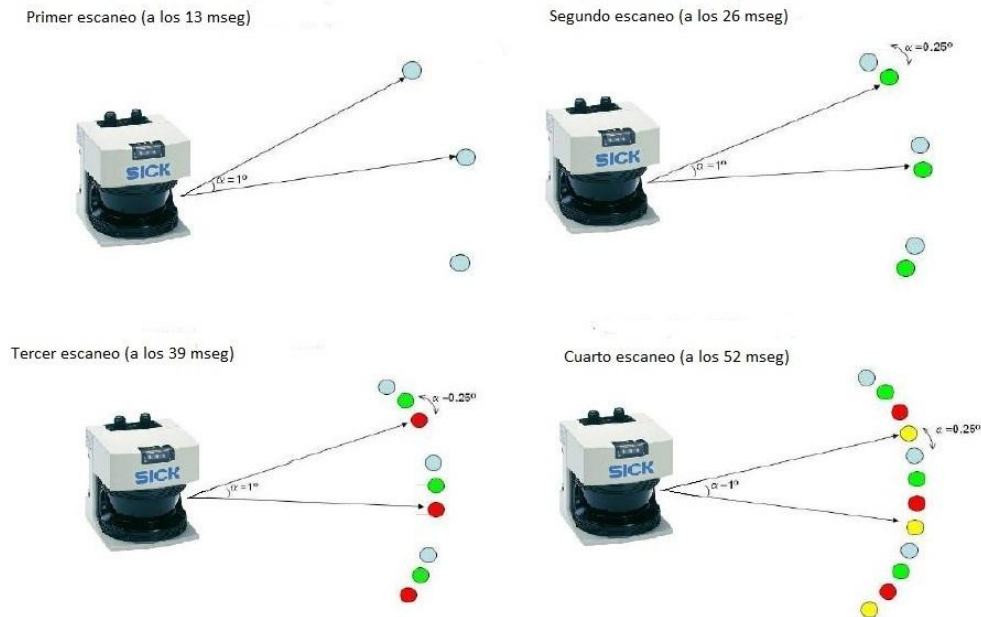


Ilustración 96. Comportamiento del láser para un escaneo de 0,25° de resolución.

4.2.2.2 Cámara de visión

La cámara de visión que vamos a utilizar es una cámara estéreo Bumblebee2 de Point Grey [19], (ilustración 97) la cual está dotada de dos objetivos. Esta cámara implementa tecnología de nueva generación que permite tiempos más rápidos en la adquisición y mejora de la calidad de datos 3D.



Ilustración 97. Cámara estéreo Bumblebee2.

La cámara estéreo Bumblebee2 incluye un sistema de calibración de segunda generación de mayor resistencia. A bordo lleva un sistema de procesamiento de color y los conectores GPIO de disparo externo además de incorporar flash.

Al ser una cámara binocular, el Bumblebee2 contiene dos 1/3-inch CCD de exploración progresiva de sensores, y transmite tanto las imágenes de la izquierda como las de la derecha a un PC mediante una interfaz IEEE-1394. Incluye las librerías Digiclops y Triclops SDK para poder trabajar con la cámara que permiten a los usuarios controlar los ajustes de la cámara, ajustar la calidad de la imagen y el acceso en tiempo real imágenes de rango de profundidad utilizando la tecnología de visión estereoscópica.

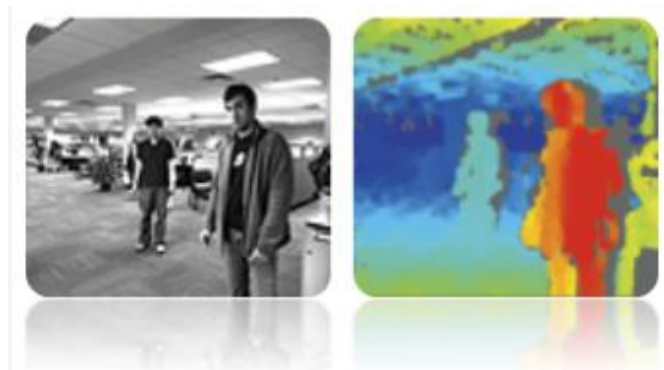


Ilustración 98. Imagen obtenida por la cámara Bumblebee2.

La cámara Bumblebee2 tiene una resolución de 640 x 480 a 48fps o 1024 x 768 a 18fps, y ofrece una opción de 3,8 o 6,0 mm de distancia focal del objetivo.

4.2.2.3 Colocación de los sensores

Ambos sensores irán colocados en el vehículo de pruebas proporcionado, el vehículo IVVI, equipado con un ordenador de a bordo y una pantalla en el salpicadero, aparte de sensores para poder medir con precisión la velocidad del vehículo y el tiempo de frenada.



Ilustración 99. Posición de los sensores en el vehículo de pruebas IVVI 2.0.

Como se puede observar en la ilustración 99 el escáner láser irá colocado en el parachoques delantero, en un soporte destinado a tal fin y la cámara estéreo de visión en el interior del vehículo, detrás del parabrisas delantero

4.2.3 Sensor GPS y sensor inercial

El MTI-G de Xsens [20] es una unidad de medición de pequeño tamaño y bajo peso, excelente para el control y la navegación en sistemas tripulados y otros objetos. El MTI-G es un GPS asistido basado en la unidad de medida inercial (IMU) y un sensor de presión estática. Ofrece un rendimiento sin precedentes por su tamaño, peso, costo y de baja complejidad en su uso. El MTI-G supera los desafíos típicos de la IMU y AHRS que se utilizan en, por ejemplo, aplicaciones aeroespaciales y de automoción.

El diseño del MTI-G es flexible, proporcionando una amplia gama de modos de salida y la configuración avanzada de escenarios de uso específico, la optimización de la fusión de sensores de rutina algoritmo para distintas aplicaciones.

El MTI-G tiene una actitud a bordo y el sistema de partida de referencia (AHRS) y el procesador de navegación. Este procesador de baja potencia de señal digital se ejecuta en tiempo real de sensores Xsens con un algoritmo de fusión, siempre sin desplazamientos, con GPS, y orientación con datos 3D. Además, el MTI-G proporciona una velocidad de datos a un ritmo de actualización más alto que con un típico receptor GPS.



Ilustración 100. Sensor GPS e inercial MTI-G de Xsens

- Aspectos más destacados:

- Cálculo en tiempo real con GPS de datos de posición y velocidad.
- La integración del GPS supera los desafíos típicos de IMU.
- AHRS integrado, GPS y sensor de presión estática.
- Sensor DSP de abordo funcionando con un algoritmo de fusión
- Frecuencia de actualización alta (120 Hz).
- Calibración individual para temperatura, desalineación en 3D y sensor de sensibilidad cruzada.
- Salida UTC referenciada.

- Facilidad de uso.

El MTI-G es un sensor que puede ser utilizado en una amplia gama de aplicaciones. Debido a los requisitos específicos para todas estas aplicaciones, el MTI-G utiliza diferentes configuraciones de filtros y restricciones, aplicadas en los

escenarios. Entre otros, hay escenarios para el uso en aplicaciones de automoción y aeroespacial.

- Salida.
 - Orientación 3D (360 °).
 - Posición y velocidad 3D (con la ayuda de sensores inerciales).
 - Aceleración 3D, tasa de giro 3D, campo magnético 3D.

4.2.4 Pantalla de visualización

El vehículo de pruebas posee una pantalla modelo Xenarc 705YV de 7 pulgadas [21], la cual se utilizará para la visualización de los resultados, la interacción con los PCs y para mostrar la señal visual de alerta si fuera necesaria.



Ilustración 101. Pantalla Xernac 705 YV de 7".



Ilustración 102. Posición de la pantalla en el interior del vehículo.

4.2.5 Comunicaciones

Tanto la cámara como el láser se comunican con un PC instalado en el interior del IVVI. Este modelo permite conectarse al láser a la máxima velocidad del puerto RS422 del láser, 500kps, y una lectura de un escaneo cada 52 milisegundos a una frecuencia de unos 19Hz.



Ilustración 103. Ordenadores de a bordo del vehículo IVVI 2.0.

4.3 Descripción del software que se va a utilizar

En esta sección se describirá el posible software que a priori se utilizará para la implementación de esta técnica de fusión sensorial.

4.3.1 Matlab

Matlab [22] (abreviatura de *MATrix LABoratory*, "laboratorio de matrices") es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M). Está disponible para las plataformas Unix, Windows y Apple Mac OS X.

Entre sus prestaciones básicas se hallan: la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con

otros dispositivos hardware. El paquete MATLAB dispone de dos herramientas adicionales que expanden sus prestaciones, a saber, Simulink (plataforma de simulación multidominio) y GUIDE (editor de interfaces de usuario - GUI). Además, se pueden ampliar las capacidades de MATLAB con las *cajas de herramientas (toolboxes)*; y las de Simulink con los *paquetes de bloques (blocksets)*.

Es un software muy usado en universidades y centros de investigación y desarrollo. En los últimos años ha aumentado el número de prestaciones, como la de programar directamente procesadores digitales de señal o crear código VHDL.

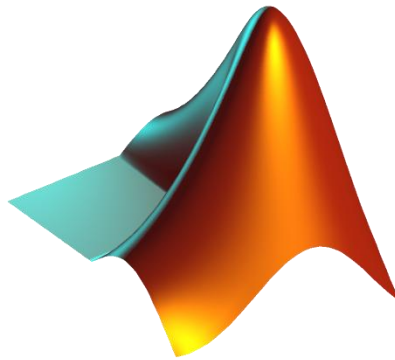


Ilustración 104. Logotipo de MATLAB

4.4 Calibración de los sensores y sincronización

Las coordenadas del sistema se pueden definir de la siguiente manera (ver ilustración 105):

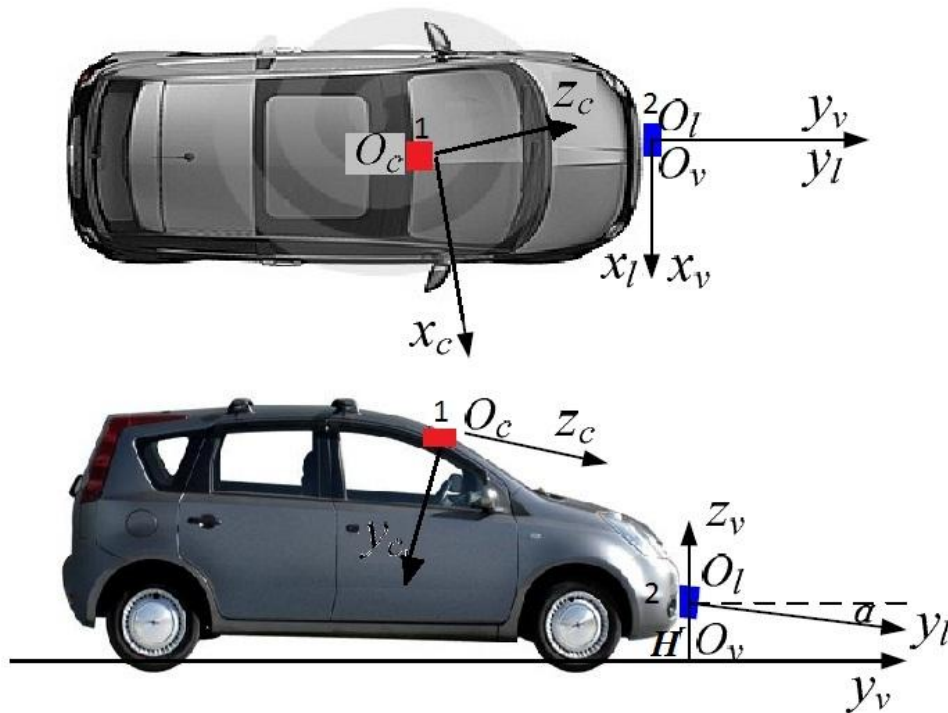


Ilustración 105. Sistemas de coordenadas de los sensores en el vehículo. 1 es la cámara y 2 el escáner láser, donde H es la altura y α el grado de inclinación del láser.

Según la ilustración 105, $O_c x_c y_c z_c$ es el sistema de coordenadas de la cámara cuyo origen es el centro óptico de la cámara. El eje z_c es el eje óptico de la cámara, los ejes $x_c y_c$ son paralelos a la dirección xy del sensor de la cámara; $O_l x_l y_l$ es el sistema de coordenadas láser cuyo origen es el centro del escáner láser. y_l es la bisectriz del ángulo del rango de exploración, y x_l es perpendicular a y_l en el plano de exploración; y $O_v x_v y_v z_v$ es el sistema de coordenadas del vehículo cuyo origen es la proyección vertical del centro de escáner láser en el suelo. x_v es paralelo a x_l .

Una vez definidos los sistemas de coordenadas, lo primero que se debe hacer es una calibración de los sensores láser y de visión, en este caso se va a utilizar un sistema de calibración similar al ya visto en [3.7.3], en el cual el escáner láser y la cámara deben integrarse en el mismo sistema de coordenadas. Para ello los puntos del láser se tienen que proyectar sobre el sistema de coordenadas de la imagen para un procesamiento posterior. La calibración tiene que incluir tanto los parámetros extrínsecos del láser y la cámara como los parámetros intrínsecos de ésta última.

Los parámetros intrínsecos son bastante estables para que la calibración intrínseca se pueda realizar con antelación. Por otro lado, el objetivo de la calibración extrínseca es encontrar los parámetros relativos de rotación y traslación entre diferentes sistemas de coordenadas.

Toda esta calibración se llevará a cabo mediante el sensor GPS e inercial MTI-G de Xsens.

Para llevar a cabo la calibración es necesario realizar una serie de transformaciones en los ejes de coordenadas de los sensores.

En primer lugar se calibran los parámetros extrínsecos del escáner láser mediante el MTI-G de Xsens. Para así trasladar las coordenadas del láser a las coordenadas del propio vehículo y a su vez al sistema de coordenadas de la cámara. Y una vez realizadas estas traslaciones las coordenadas del vehículo pueden ser proyectadas en el sistema de coordenadas de la imagen.

4.5 Adquisición de datos

Para esta parte se tomarán los datos obtenidos por el escáner láser por un lado y la cámara de visión por otro, y una vez adquiridos procesarlos mediante un algoritmo de fusión.

4.5.1 Sistema del escáner láser

Los datos proporcionados por el escáner láser serán un número de puntos fijos que deben ser agrupados en objetos, estos objetos se pueden clasificar en cuatro grupos distintos y además aquellos puntos que no siguen ningún patrón o su distancia entre sí es superior al rango establecido se descartarán. Una vez que los puntos se han agrupado en segmentos llamados polilíneas que tienen una forma característica para

cada tipo de obstáculo. Los grupos en los que se pueden clasificar los objetos detectados serán:

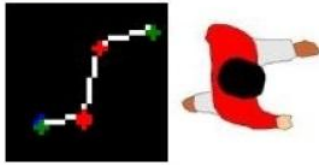
- a) Posibles peatones.
- b) Infraestructura de la carretera tales como bordes de la carretera, guardarraíles...
- c) Obstáculos en forma de L, correspondientes a vehículos estacionados.
- d) Otros obstáculos.

Los objetos que son detectados como no peatones, es decir, infraestructura de la carretera, obstáculos en forma de L u otros obstáculos se descartarán quedándonos únicamente con los posibles peatones.

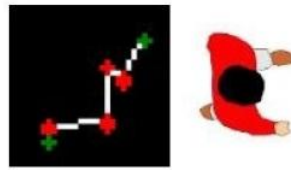
Los obstáculos detectados como posibles peatones tendrán una forma característica según los patrones explicados en [3.2.2.2] y así detectar los peatones que caminan cruzando un paso de peatones (perpendicular a la dirección del vehículo) o alejándose/acercándose al vehículo ya sea paralelamente u oblicuamente.

En las siguientes ilustraciones se pueden observar dos secuencias de peatones caminando perpendicularmente y alejándose del láser, respectivamente:

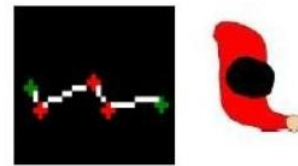
El segmento superior corresponde a la pierna izquierda y el segmento inferior corresponde a la pierna derecha.



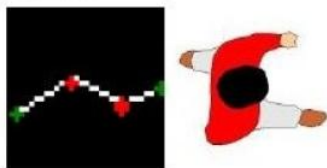
La pierna izquierda permanece quieta, mientras el pie derecho está avanzando.



El peatón tiene las dos piernas casi juntas, por lo que no se diferencia la izquierda de la derecha.



El peatón tiene la pierna derecha en una posición mas avanzada que la pierna izquierda, ha completado un paso.



La pierna izquierda avanza para volver a empezar un paso.

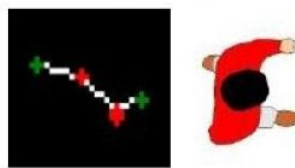
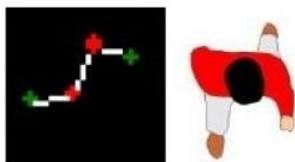
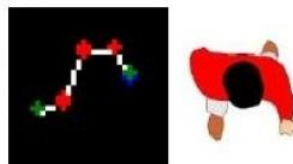


Ilustración 106. Secuencia de un peatón caminando perpendicularmente al láser.

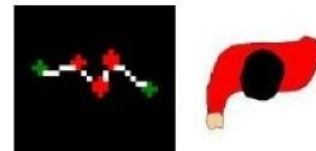
El peatón tiene la pierna derecha más adelantada que su pierna izquierda.



El peatón está llevando la pierna izquierda a la altura de su pierna derecha.



El peatón tiene las dos piernas a la misma altura, el pie derecho esta apoyado en el suelo, y el izquierdo en el aire con movimiento hacia delante.



El peatón tiene la pierna izquierda por delante de la derecha, sin apoyar el pie derecho en el suelo.



El pie izquierdo esta apoyado en el suelo, el peatón ha completado un paso.

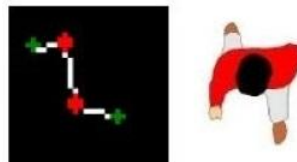


Ilustración 107. Secuencia de un peatón acercándose al láser.

- Los obstáculos detectados como infraestructura de la carretera serán obstáculos de grandes dimensiones situados en los laterales de la vía.

- Los obstáculos en forma de L tendrán dos líneas que formarán un ángulo recto, correspondientes a un lateral y al frontal o a la parte posterior de un vehículo estacionado.
- Los obstáculos que formen polilíneas y que no estén incluidos en ninguno de los grupos anteriores se clasificarán como “otros obstáculos”.

El diagrama correspondiente al algoritmo de clasificación para los posibles peatones se puede observar en la ilustración 108.



Ilustración 108. Diagrama de flujo del algoritmo de clasificación de peatones.

4.5.2. Sistema de visión en estéreo

Por su parte, la cámara de visión capturará los fotogramas correspondientes al movimiento de los peatones de la región de interés delimitada por el escáner láser, y mediante un algoritmo de fusión de los datos obtenidos previamente por el láser y la cámara se verificará si la detección por el escáner láser del posible peatón es correcta.

Una de las diferencias fundamentales entre un sistema de detección estéreo y aquellos compuestos de una única cámara es que el sistema de visión estereoscópica permite determinar no sólo la lejanía a la que se encuentra el objeto detectado con respecto a nuestro vehículo sino que además permite determinar la altura a la que dicho objeto se encuentra sobre la superficie del asfalto, función que no pueden realizar las mono-cámaras, ya que éstas únicamente pueden estimar distancias con respecto a nuestro vehículo.

Además, la doble cámara permite que al analizar las diferencias entre las dos imágenes tomadas, el sistema sea capaz de establecer de qué tipo de obstáculo se trata (desde posibles mercancías que han caído sobre la calzada, peatones o animales), por lo que se avanza en el desarrollo de sistemas de detección de peatones. Esta diferenciación no puede hacerse con la suficiente fiabilidad con los sistemas mono-cámara, que han sido diseñados para reconocer coches o motocicletas y, por lo tanto, son sólo capaces de identificar objetos que previamente han aprendido.

Mediante la combinación de las imágenes obtenidas en ambas cámaras y estableciendo qué aspectos entre ellas resultan iguales y cuáles son diferentes, el nuevo sistema de detección es capaz de reconocer los objetos incluso en aquellas condiciones donde otros sistemas se encuentran limitados, por ejemplo cuando varios objetos se encuentran muy cercanos entre ellos, objetos poco o mal iluminados o cuando entre el objeto y su entorno hay poco contraste.

El análisis píxel a píxel entre cada uno de los fotogramas que recoge la cámara permite establecer el patrón de movimiento del peatón, lo que permite calcular si la trayectoria de dicho peatón interferirá en la trayectoria que lleva el vehículo.

Una vez recogidos los fotogramas mediante un algoritmo se analizarán para detectar los posibles peatones que haya en la imagen. El algoritmo también ha de ser capaz de detectar peatones parcialmente ocluidos.



Ilustración 109. Ejemplo de imágenes típicas que pueden ser obtenidos por la cámara, la primera fila muestra peatones, la segunda fila imágenes en las que no hay peatones y la tercera fila peatones parcialmente ocluidos.

El algoritmo para detectar los posibles peatones tiene que obtener una lista de regiones en la imagen conteniendo a peatones potenciales, y para ello se explotarán los rasgos propios de los humanos: en concreto, los bordes verticales y la simetría vertical que presenta su silueta.

Este enfoque se basa en el hecho que el contorno de un peatón presenta más bordes verticales que horizontales. El uso de la cámara estéreo del espectro visible, permite capturar mayor nivel de detalles que las cámaras de infrarrojos. Esto implica que las imágenes contendrán más bordes, pero también mayor cantidad de ruido, siendo necesario realizar un preprocesamiento previo a la extracción de bordes.

Una vez detectados los bordes verticales se realiza la búsqueda de simetrías verticales en cada una de esas imágenes. Se recorre toda la imagen, pero el número de puntos que se deben considerar son los pertenecientes a los bordes verticales, resultando rápido. El tamaño de la ventana de búsqueda queda establecido en función del rango de disparidades contenido en cada imagen.

De este modo, se obliga a que la búsqueda de simetrías cumpla unas restricciones en cuanto al tamaño estimado del peatón.

A partir de la búsqueda de simetrías se determinan las regiones de interés. Para tomar esta decisión, sólo se considerarán aquellas ventanas que cumplan las condiciones geométricas antes citadas. Pero, debido a que en el entorno puede haber otros objetos con una simetría vertical fuerte, es necesario filtrar las regiones candidatas.

Se han impuesto restricciones en cuanto al ancho mínimo que deben tener los bordes considerados para obtener el eje de simetría. De este modo se eliminan objetos como postes, farolas o árboles, que son elementos habituales en entornos urbanos y que ocasionan multitud de falsos positivos. Además, se exige que el número de votos de cada ventana esté por encima de un umbral, que varía en función del tamaño de la ventana en evaluación. Así, para la selección del mejor candidato se considera la relación existente entre, el número de puntos de borde simétricos y el tamaño de la ROI que los contiene.

Una vez detectado el posible peatón se procederá a validarlo mediante el algoritmo de fusión.

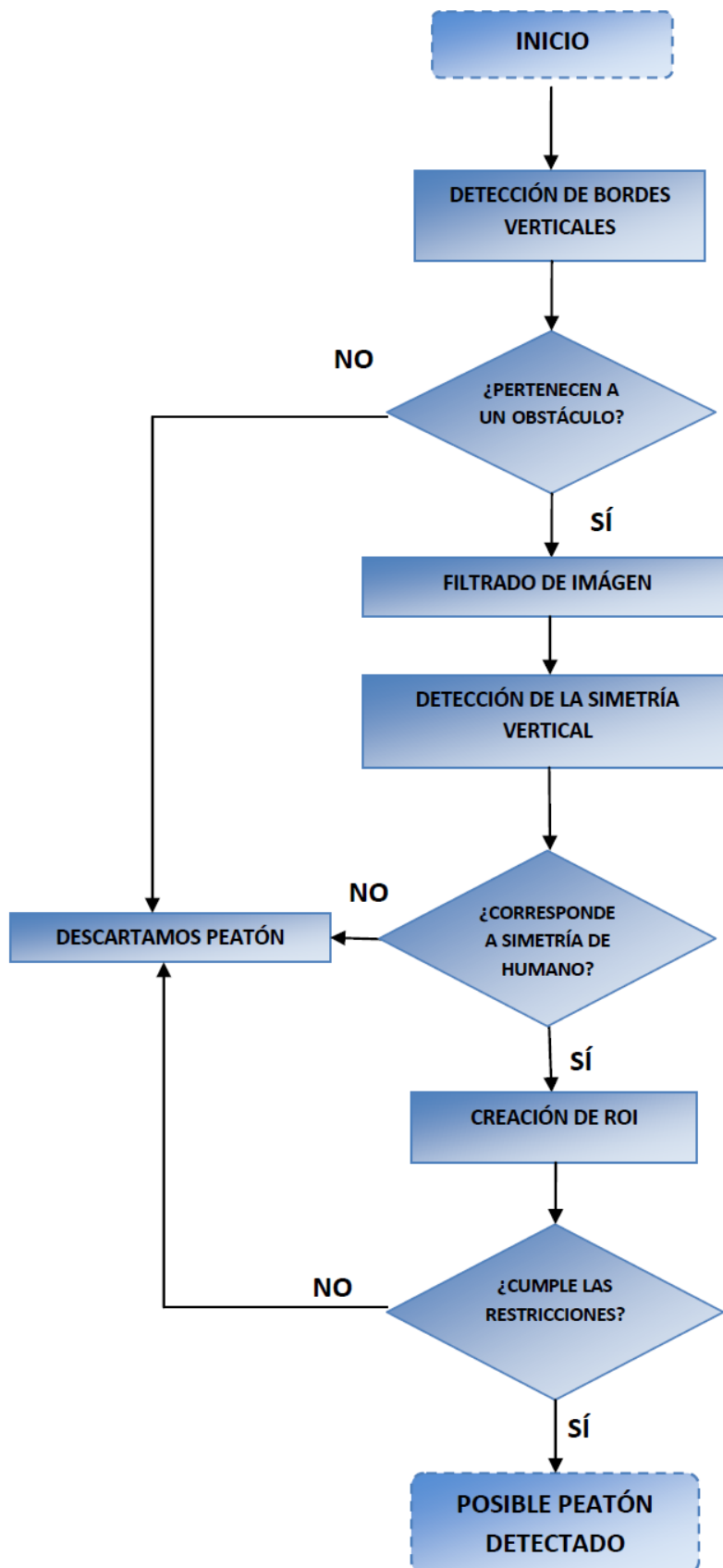


Ilustración 110. Diagrama de flujo para algoritmo del subsistema de visión.

4.5.3. Fusión de datos.

Una vez adquiridos los datos de ambos sensores y dado que se han tenido que calibrar previamente, únicamente se utiliza la visión para validar el candidato generado por el láser y se clasifica según lo visto anteriormente en 4.5.1.

Como el telémetro láser y la cámara estéreo tienen diferentes rangos de captura se eliminarán todos los datos que no están comprendidos en ambos rangos. Por otro lado sólo se analizarán aquellos datos que se hayan obtenido al mismo tiempo por el escáner láser y la cámara estéreo, en caso contrario se desecharían.

Para comenzar se analizarían los datos obtenidos por el escáner láser en busca de un posible peatón para un tiempo concreto (un frame). Si el algoritmo del escáner láser detecta un posible peatón se procederá a validarlo mediante los datos obtenidos por la cámara estéreo de visión en ese mismo frame. Si el algoritmo de la cámara detecta el mismo peatón o peatones que el láser en ese frame, se cumpliría que efectivamente se ha detectado un peatón y se procedería a realizar el seguimiento de su trayectoria, si por el contrario en los datos obtenidos por la cámara de visión no se detecta ningún peatón o no coincide con el detectado mediante el escáner láser se consideraría un falso positivo y se desecharía.

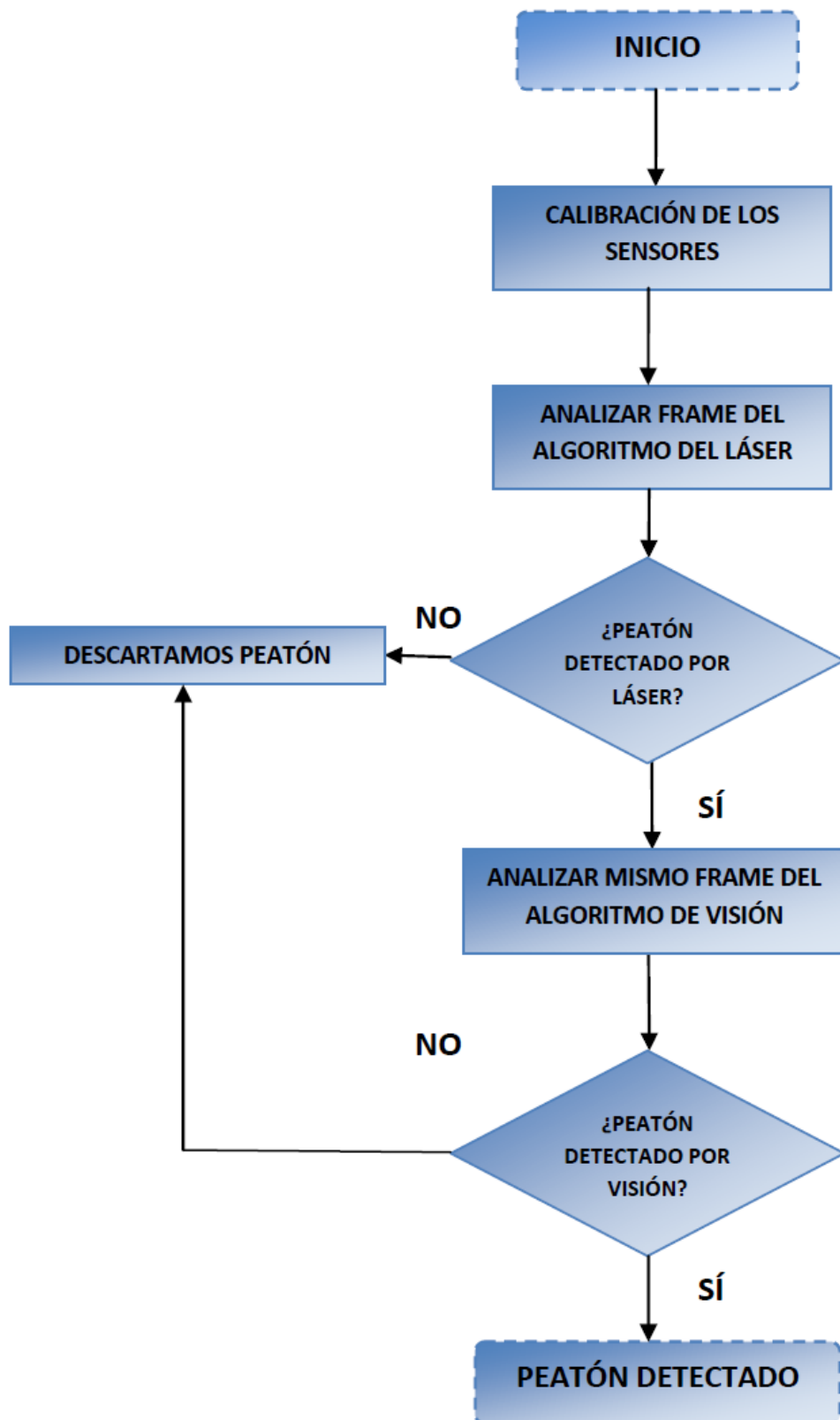


Ilustración 111. Diagrama de flujo del algoritmo de fusión.

4.5.4. Seguimiento.

Una vez fusionados los datos se tiene que proceder al seguimiento de su trayectoria, para así predecir los posibles movimientos de los peatones y anticiparse a la posible colisión.

Para ello nos vamos a servir de cuatro algoritmos distintos: el filtro de Kalman [24], el filtro de Kalman Extendido [25], el filtro de Kalman Unscented [26] y el filtro de Partículas [27]. Con esto pretendemos conseguir realizar de cuatro maneras distintas el seguimiento de la trayectoria del peatón, y así localizar los falsos positivos y los errores que se produzcan en la detección.

4.6 Sistema de alerta

El sistema de frenada de emergencia que se va a describir a continuación es un sistema teórico que podía implementarse en el futuro, actualmente es inviable colocarlo en el vehículo de pruebas IVVI por causas de la legislación de tráfico, habría que modificar el vehículo para poder circular por la vía pública, lo que supondría un aumento elevado del coste.

Para diseñar el sistema de alerta nos basaremos en [3.3]. El algoritmo de alerta proporciona una representación visual y un sistema de alerta auditiva en tiempo real al conductor en función del riesgo de colisión, además si el tiempo estimado de colisión es lo suficientemente bajo para que exista riesgo real de atropello el sistema de frenada automática se tiene que poner en marcha. El algoritmo de advertencia estima que el riesgo de colisión de la posición predicha del peatón mediante un sistema de tiempo de colisión (T_{TC}) que se tendrá que medir de acuerdo a un algoritmo que relacione la velocidad a la que va el vehículo, la distancia de la posible colisión y el tiempo de reacción del conductor. Una vez estimado el tiempo de colisión se clasificará el riesgo de colisión según un umbral. Si el riesgo es moderado se



efectuará un aviso auditivo y si el riesgo de colisión es inminente se efectuaría una frenada de emergencia.

El diagrama de flujo del algoritmo de alerta se muestra en la ilustración 112.

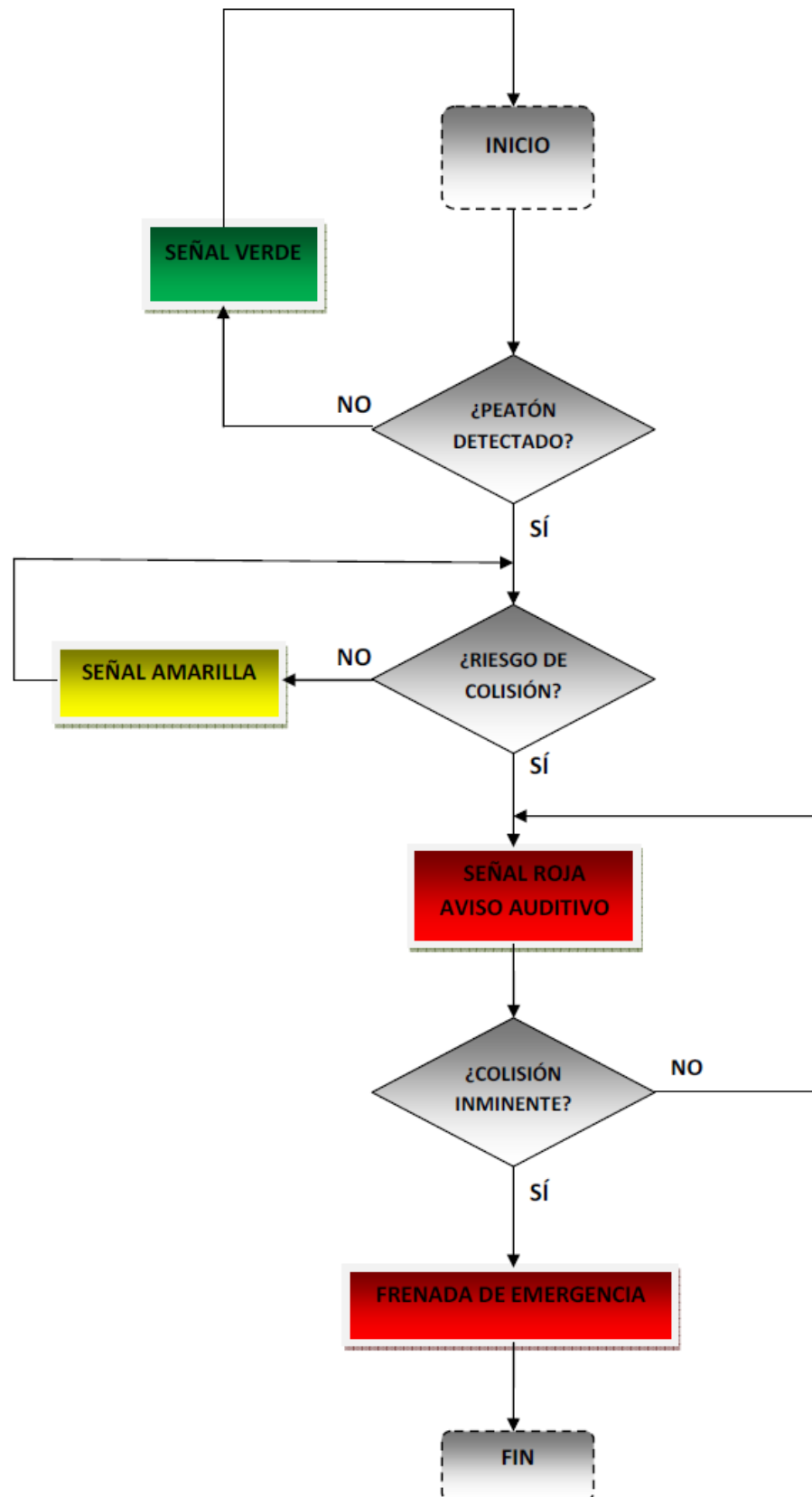


Ilustración 112. Diagrama de flujo del algoritmo de alerta.

Lo primero será definir la dirección del movimiento del peatón, y después calcular el tiempo previsto (T_p) del vehículo para llegar a dicha posición. La posición predicha para los peatones es la posición en la que un peatón se mueve por el tiempo (T_p). No obstante se supone que tanto el peatón como el vehículo se mueven a una velocidad constante ($V_{ped} = cte$, $V_{car} = cte$). El vehículo de pruebas se moverá en línea recta y perpendicular al paso de peatones ($Y_{car} = 0$). Los peatones se moverán a lo largo del eje móvil definido anteriormente con una velocidad media constante ($X_{ped} = cte$). Una vez hecho esto, mediante las ecuaciones 3 y 4 [3.3] se calculará la distancia del vehículo al peatón y el grado de peligrosidad.

Después, para comprobar los datos de salida de procesamiento del sistema, como se puede observar en la ilustración 113, se utilizará un semáforo en rojo / amarillo / verde dentro del vehículo:



Ilustración 113. Información presentada en la pantalla

Dependiendo de la condición de la posición de los peatones y su velocidad al caminar, la advertencia de colisión consiste en una señal auditiva, y además el color de la advertencia visual en la pantalla del vehículo cambia de acuerdo a las siguientes reglas, y se expedirá un aviso auditivo al conductor cuando la situación se vuelva más crítica, llegando incluso (si el vehículo se adaptara convenientemente

según la legislación de tráfico) a realizar una frenada de emergencia automática si hay riesgo inminente de atropello:

- d) Si no se detecta ni paso de peatones ni ningún peatón, el color mostrado en la pantalla será verde.
- e) Si se detecta un paso de peatones o peatones cerca del vehículo el color mostrado en la pantalla será amarillo y además se mostrará en la pantalla "Paso de peatones detectado".
- f) Si se detecta un paso de peatones y peatones en movimiento en el paso de peatones y la posición predicha está por debajo del valor de los valores de T_{TC} calculado, el color mostrado por la pantalla será rojo y se mostrará en la pantalla, "Paso de peatones ocupado" y además si el valor de T_{TC} establece un riesgo elevado de colisión, se proporciona al conductor un aviso auditivo.
- g) Si además de darse las condiciones anteriores, el valor de T_{TC} estima que la colisión es inminente se activaría el sistema de frenada de emergencia.

En la ilustración 114 se puede ver gráficamente lo explicado anteriormente.

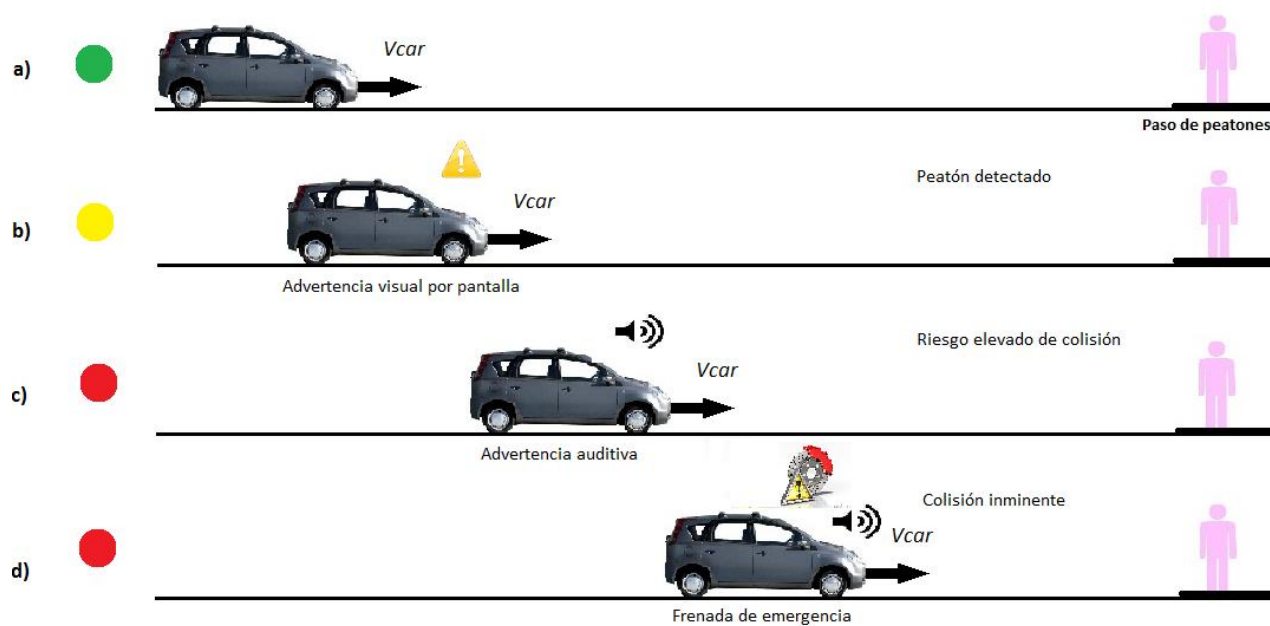


Ilustración 114. Descripción gráfica del sistema de alerta y frenada de emergencia.

Capítulo 5. Resultados

5.1. Introducción

Se van a analizar 15 secuencias correspondientes a 8 peatones para realizar el seguimiento de sus trayectorias. Como se va a ver las dos primeras secuencias correspondientes al peatón 1 se han analizado mediante un filtro de Kalman, un filtro de Kalman Unscented y un filtro de partículas. Las demás secuencias se han analizado únicamente mediante el filtro de Kalman debido a las similitudes encontradas con los demás filtros de estimación. Las secuencias fueron grabadas desde el vehículo de pruebas IVVI 2.0 e incluyen la información tanto del láser, la cámara de visión estéreo y el sistema de fusión en conjunto.

En las pruebas participaron 8 peatones en total que caminaban paralelamente a una distancia respecto al vehículo de 5 y 10 metros haciendo un recorrido de ida y vuelta con lo que tenemos 15 secuencias distintas (el peatón número 8 sólo hizo el recorrido a 10 metros). Los frames en los que el peatón se sale del rango de visión de la cámara así como los frames en los que se produjo una falta de sincronización no se han tenido en cuenta a la hora de los cálculos.

De las pruebas se han obtenido el número de positivos detectados tanto del láser, de la cámara y del sistema de fusión en conjunto así como los porcentajes respecto a los frames totales. Y además se han obtenido los falsos positivos que aparecieron en el láser, la cámara y el sistema de fusión. Todas estas estimaciones se han realizado con el filtro de Kalman. Las estimaciones realizadas mediante el filtro de Kalman Unscented y el filtro de partículas dan unos resultados muy similares y sólo se han obtenido de ellos las gráficas pertinentes y errores.



5.2 Secuencias analizadas mediante el filtro de Kalman

A continuación se van a analizar dos secuencias de los peatones mediante el algoritmo del filtro de Kalman para obtener los positivos, los falsos positivos y el error medio y total de la estimación para el seguimiento del peatón. Para que no resulte un capítulo demasiado amplio para el resto de secuencias se mostrará únicamente una tabla con los datos finales.

Significado de los colores en cada gráfica:

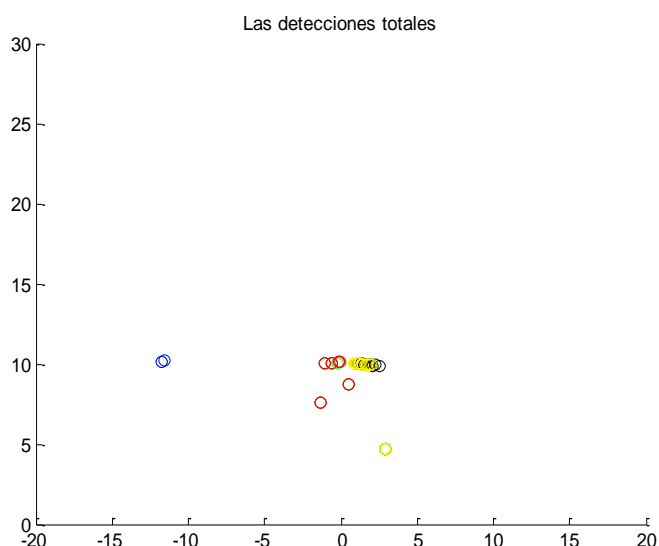
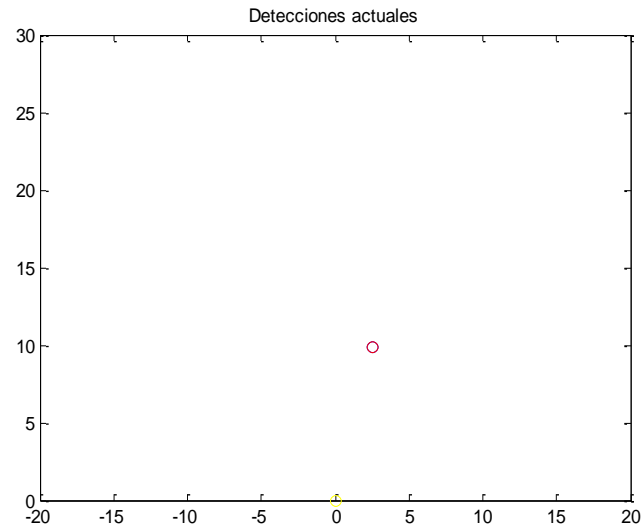


Ilustración 115. Detecciones totales

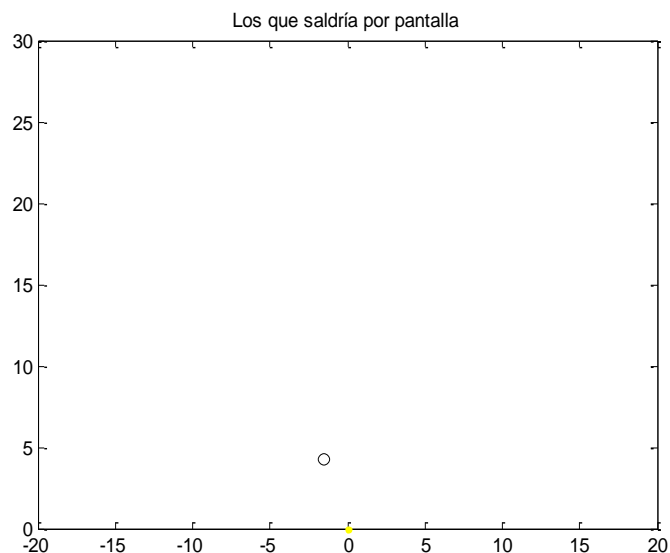
Esta gráfica representa todas y cada una de las detecciones de los peatones en los dos sensores, tanto del laser como de la cámara, con los siguientes significados:

- Amarillo: las estimaciones hechas con el filtro de Kalman para detecciones consolidadas
- Negro: Valor de la detección de una detección consolidada (actualización)
- Rojo: Detección no consolidada para visión (no ha habido detección del laser)
- Azul: Detección no consolidada para laser (no ha habido detección de la cámara)
- Verde: Las detecciones que no estamos haciendo el seguimiento (Nuevas detecciones)

**Ilustración 116. Detecciones actuales**

En esta gráfica se representan las detecciones de los sensores en un instante determinado.

- Azul: Información de la cámara disponible. Detección únicamente del laser.
- Rojo: Detección de la cámara y del láser.
- Verde: Detección únicamente de la cámara.
- Negro: Información de la cámara no disponible. Detección del laser.
-

**Ilustración 117. Lo que saldría por pantalla**

Esta figura es la hipotética salida del sistema de fusión en un instante determinado:

- Negro: Detección actualizada
- Verde: Detección no actualizada

Ambas siguen contando para la detección total.

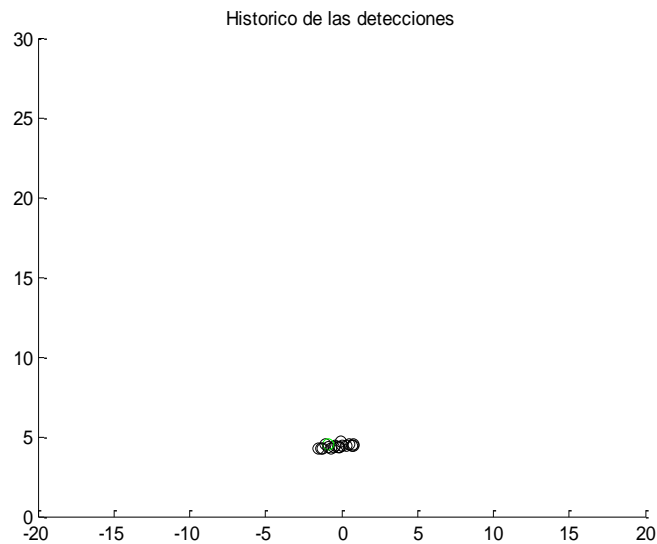


Ilustración 118. Histórico de las detecciones.

Aquí, igual que en la gráfica anterior es la salida del sistema de fusión, pero en vez de en un instante determinado, se va guardando un histórico de las detecciones del sistema, tanto actualizadas como no actualizadas.

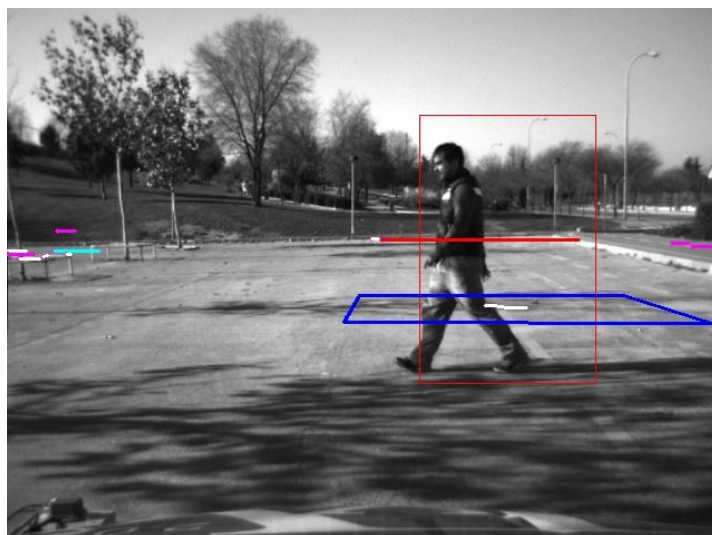


Ilustración 119. Frame de la secuencia.

En esta imagen se puede observar las detecciones del láser y de la cámara.

- Rectángulo rojo: detección de la cámara.
- Paralelepípedo azul: detección del láser.

5.2.1. Peatón 1

5.2.1.1 Secuencia 1

A 5 metros estos son los resultados obtenidos por el filtro de Kalman:

Nº de frames totales: 26

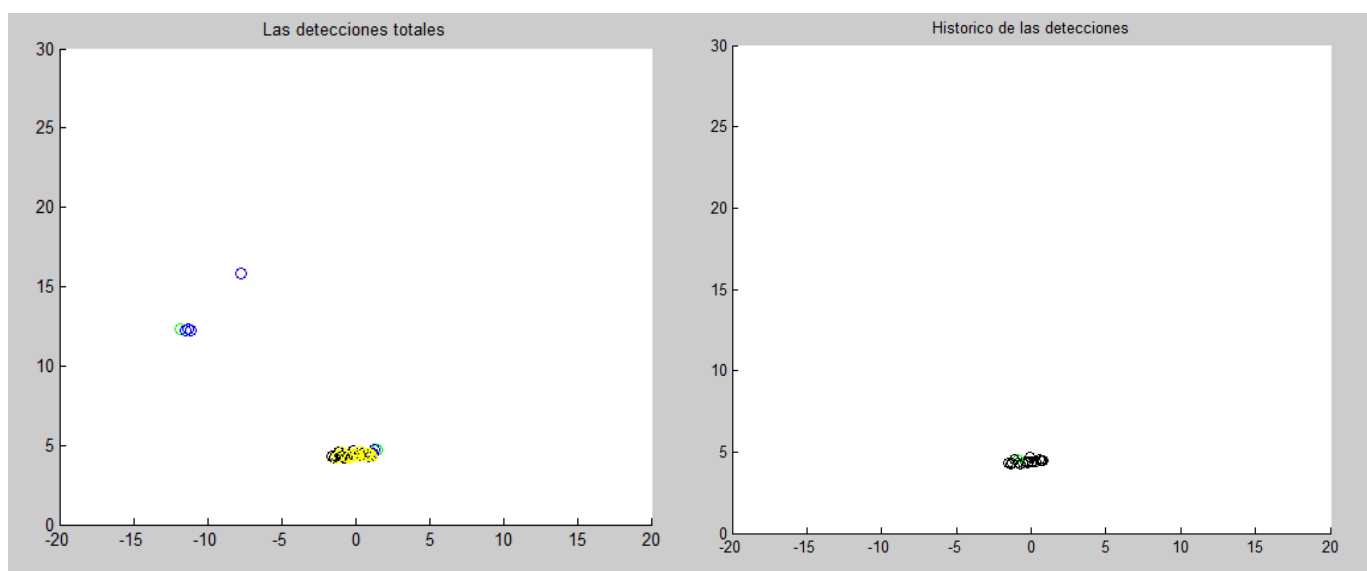
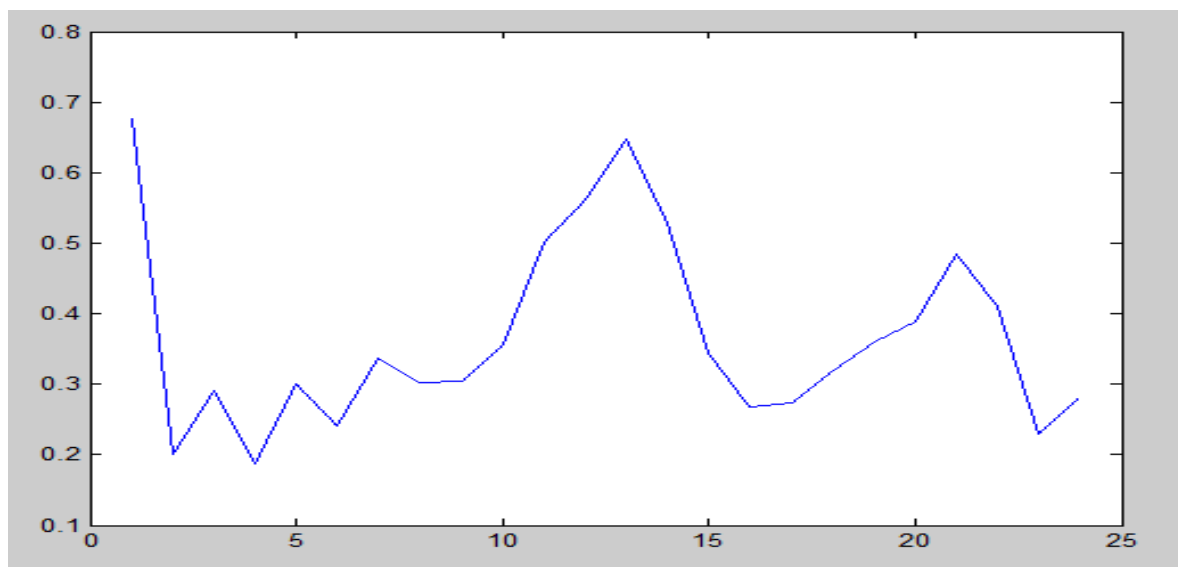
Nº de frames de la cámara (información disponible de la cámara): 14

Tabla 5. Secuencia 1

| | Láser | Cámara | Fusión |
|--------------------|---------------|---------------|---------------|
| Positivos | 24 | 6 | 19 |
| % Positivos | 92.3% | 42.85% | 73.07% |
| Falsos positivos | 3 | 0 | 0 |
| % Falsos positivos | 11.53% | 0% | 0% |

Error total de la secuencia: 0,3664

Error por frame:



5.2.1.2 Secuencia 2

A 10 metros estos son los resultados obtenidos por el filtro de Kalman:

Nº de frames totales: 102

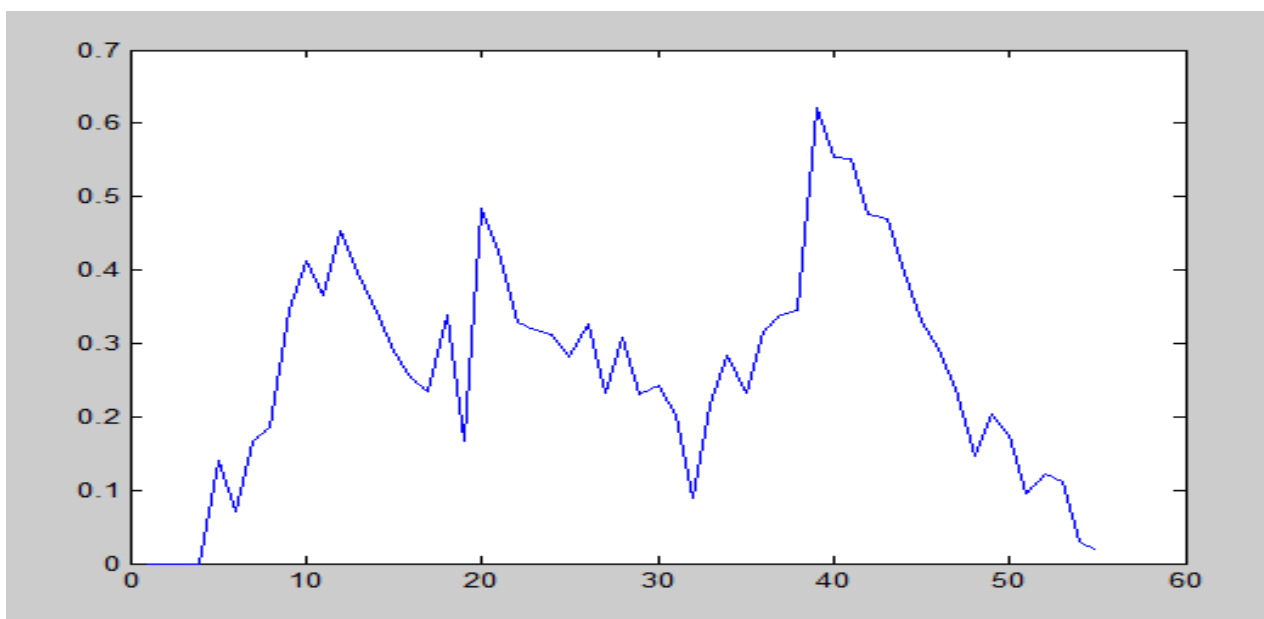
Nº de frames de la cámara (información disponible de la cámara): 54

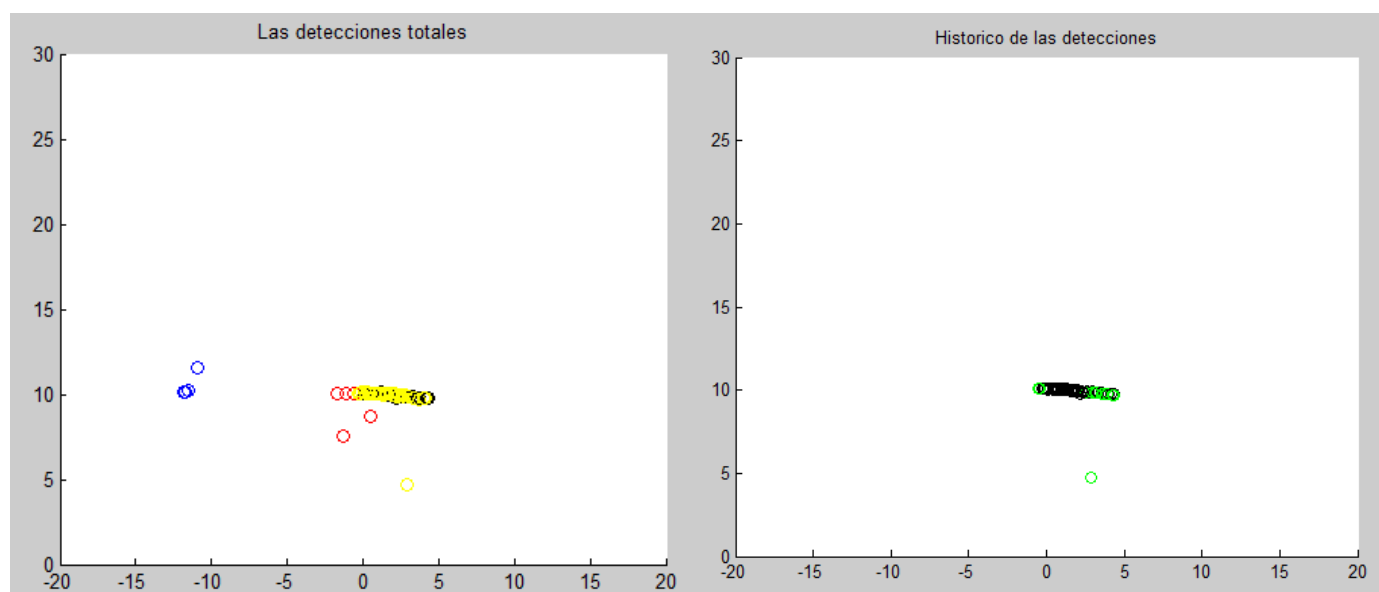
Tabla 6. Secuencia 2

| | Láser | Cámara | Fusión |
|--------------------|---------------|---------------|---------------|
| Positivos | 42 | 24 | 59 |
| % Positivos | 41.17% | 44.44% | 57.84% |
| Falsos positivos | 8 | 1 | 7 |
| % Falsos positivos | 7.84% | 1.85% | 6.86% |

Error total de la secuencia: 0,2635

Error por frame:





5.2.2 Resto de peatones

Estos son los resultados de las 15 secuencias correspondientes a los ocho peatones a los que se les ha realizado el seguimiento. Las secuencias 3 y 4 corresponden al peatón 2, las 5 y 6 al peatón 3 y así sucesivamente hasta el peatón número 8 que sólo tendrá la secuencia a 10 metros de distancia:

Tabla 7. Resto de secuencias analizadas

| <i>Resto de secuencias analizadas</i> | | | |
|---------------------------------------|--------|--------|--------|
| Secuencia 3 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 49 | 37 | 49 |
| % Positivos | 81.66% | 94.87% | 81,66% |
| Falsos positivos | 3 | 3 | 0 |
| % Falsos positivos | 5% | 7.69% | 0% |
| Secuencia 4 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 77 | 43 | 81 |
| % Positivos | 79.3% | 63.2% | 83.5% |
| Falsos positivos | 10 | 4 | 5 |
| % Falsos positivos | 10.3% | 5.88% | 5.15% |
| Secuencia 5 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 65 | 47 | 65 |
| % Positivos | 90.27% | 79.6% | 90.27% |
| Falsos positivos | 0 | 10 | 2 |

| | | | |
|---------------------|--------|--------|--------|
| % Falsos positivos | 0% | 16.94% | 2.77% |
| Secuencia 6 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 87 | 51 | 93 |
| % Positivos | 91.5% | 71.8% | 97.8% |
| Falsos positivos | 1 | 4 | 4 |
| % Falsos positivos | 1.05% | 5.63% | 4.21% |
| Secuencia 7 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 73 | 55 | 71 |
| % Positivos | 90.1% | 94.49% | 87.65% |
| Falsos positivos | 3 | 7 | 2 |
| % Falsos positivos | 3.7% | 12.28% | 2.46% |
| Secuencia 8 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 54 | 43 | 52 |
| % Positivos | 85.7% | 89.5% | 82.5% |
| Falsos positivos | 1 | 4 | 1 |
| % Falsos positivos | 1.58% | 8.33% | 1.58% |
| Secuencia 9 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 92 | 43 | 90 |
| % Positivos | 97.8% | 71.6% | 95.7% |
| Falsos positivos | 10 | 3 | 0 |
| % Falsos positivos | 10.63% | 5% | 0% |
| Secuencia 10 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 82 | 46 | 83 |
| % Positivos | 70.08% | 61.3% | 70.9% |
| Falsos positivos | 4 | 0 | 0 |
| % Falsos positivos | 3.41% | 0% | 0% |
| Secuencia 11 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 48 | 31 | 49 |
| % Positivos | 75% | 68.8% | 76.2% |
| Falsos positivos | 4 | 5 | 0 |
| % Falsos positivos | 6.25% | 11.11% | 0% |
| Secuencia 12 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 49 | 29 | 54 |
| % Positivos | 44.5% | 42.69% | 49.09% |
| Falsos positivos | 4 | 1 | 0 |
| % Falsos positivos | 3.63% | 1.47% | 0% |
| Secuencia 13 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 65 | 27 | 61 |
| % Positivos | 100% | 75% | 93.8% |
| Falsos positivos | 3 | 3 | 0 |

| | | | |
|---------------------|-------|--------|--------|
| % Falsos positivos | 3.63% | 1.47% | 0% |
| Secuencia 14 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 79 | 44 | 85 |
| % Positivos | 63.7% | 60.2% | 68.5% |
| Falsos positivos | 11 | 1 | 1 |
| % Falsos positivos | 8.87% | 1.36% | 0.8% |
| Secuencia 15 | | | |
| | Láser | Cámara | Fusión |
| Positivos | 81 | 33 | 85 |
| % Positivos | 95.2% | 66% | 100% |
| Falsos positivos | 10 | 3 | 1 |
| % Falsos positivos | 8.87% | 1.36% | 0.8% |

5.3 Secuencias por Filtro de Kalman Unscented y Filtro de Partículas

En este apartado se muestran las gráficas de las detecciones totales y el histórico de las detecciones pero esta vez realizado mediante dos estimadores distintos, el Filtro de Kalman Unscented y el Filtro de Partículas. El peatón 1 ha sido analizado de la misma manera que en el apartado 5.2 calculando los positivos, los falsos positivos y el error total. Dada la similitud de los resultados obtenidos mediante los distintos estimadores para realizar el seguimiento de la trayectoria, en los demás peatones se ha omitido este cálculo.

5.3.1 Peatón 1 analizado por Filtro de Kalman Unscented

5.3.1.1 secuencia 1

A 5 metros estos son los resultados obtenidos por el filtro de Kalman Unscented:

Nº de frames totales: 26

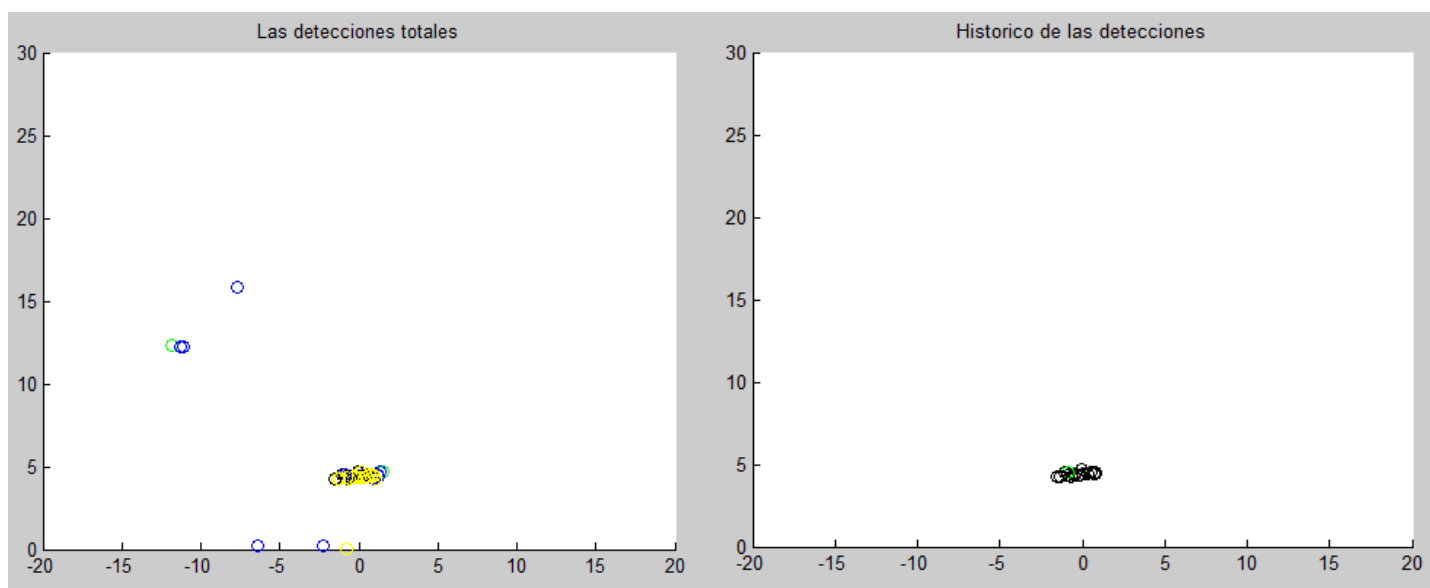
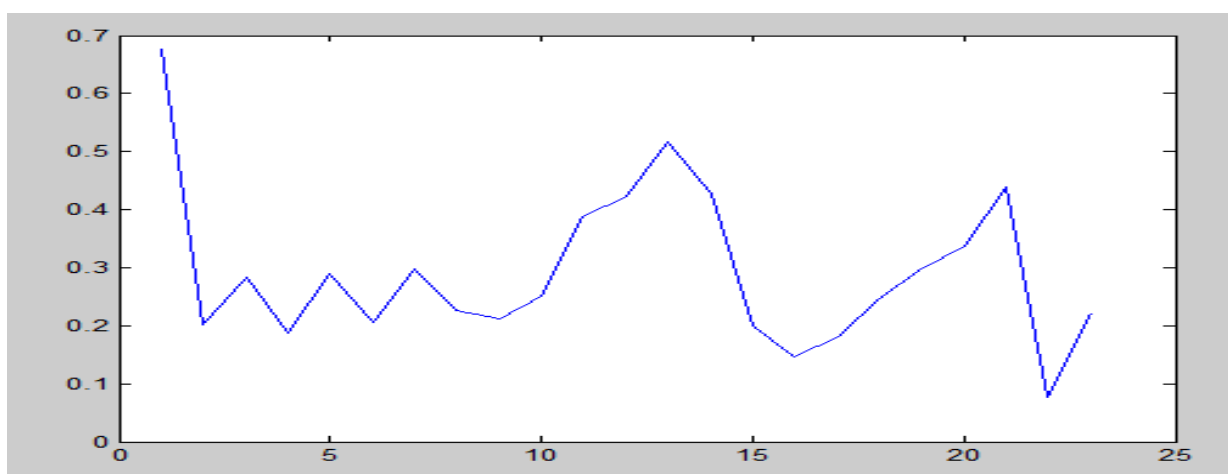
Nº de frames de la cámara (información disponible de la cámara): 14

Tabla 8. Secuencia 1UKF

| | Láser | Cámara | Fusión |
|--------------------|--------|--------|--------|
| Positivos | 25 | 6 | 18 |
| % Positivos | 96.15% | 42.85% | 69.23% |
| Falsos positivos | 3 | 0 | 0 |
| % Falsos positivos | 11.53% | 1.36% | 0.8% |

Error total de la secuencia: 0,2926

Error por frame:



5.3.1.2 Secuencia 2

A 10 metros estos son los resultados obtenidos por el filtro de Kalman Unscented:

Nº de frames totales: 102

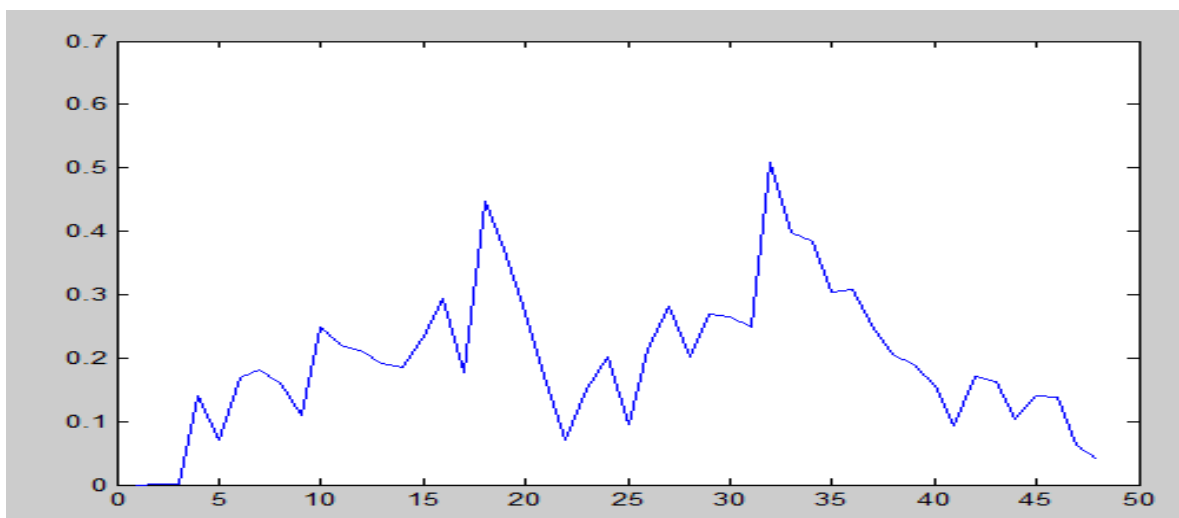
Nº de frames de la cámara (información disponible de la cámara): 54

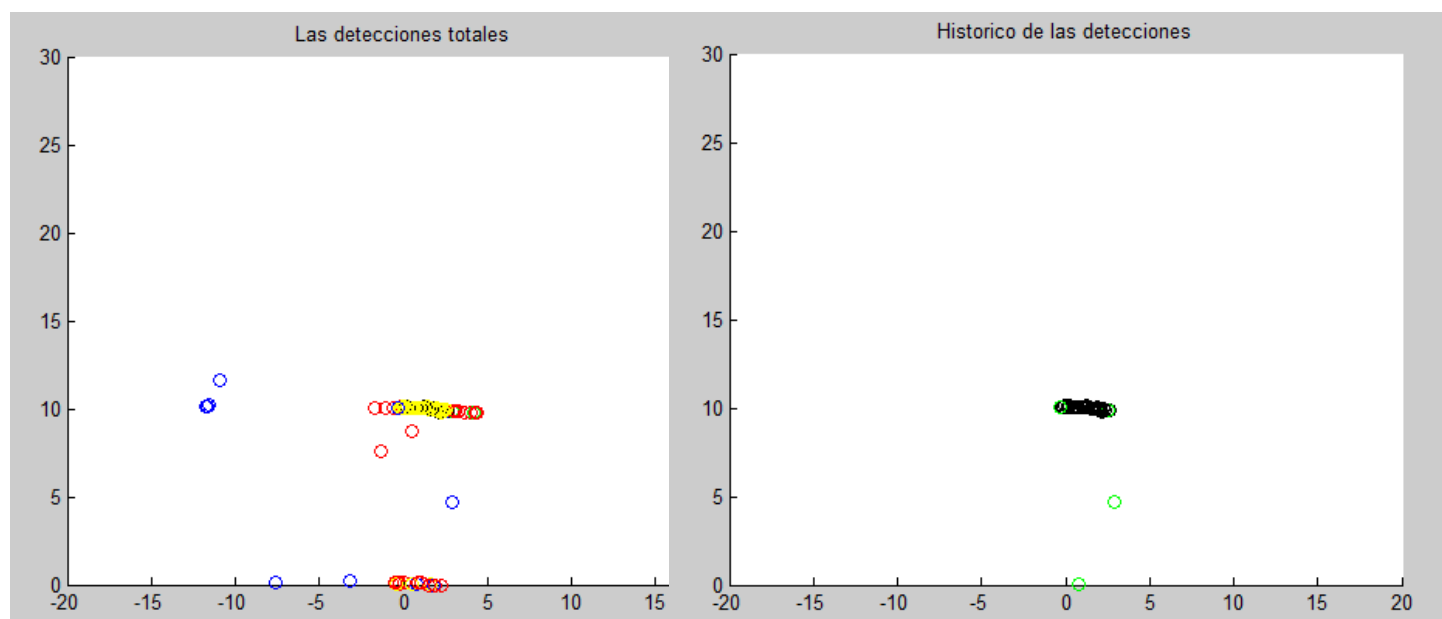
Tabla 9. Secuencia 2 UKF

| | Láser | Cámara | Fusión |
|--------------------|---------------|---------------|---------------|
| Positivos | 44 | 24 | 43 |
| % Positivos | 43.13% | 44.44% | 42.15% |
| Falsos positivos | 7 | 1 | 4 |
| % Falsos positivos | 6.86% | 1.85% | 3.92% |

Error total de la secuencia: 0,1969

Error por frame:





5.3.2 Peatón 1 analizado por Filtro de Partículas

5.3.2.1 Secuencia 1

A 5 metros estos son los resultados obtenidos por el filtro de Partículas:

Nº de frames totales: 26

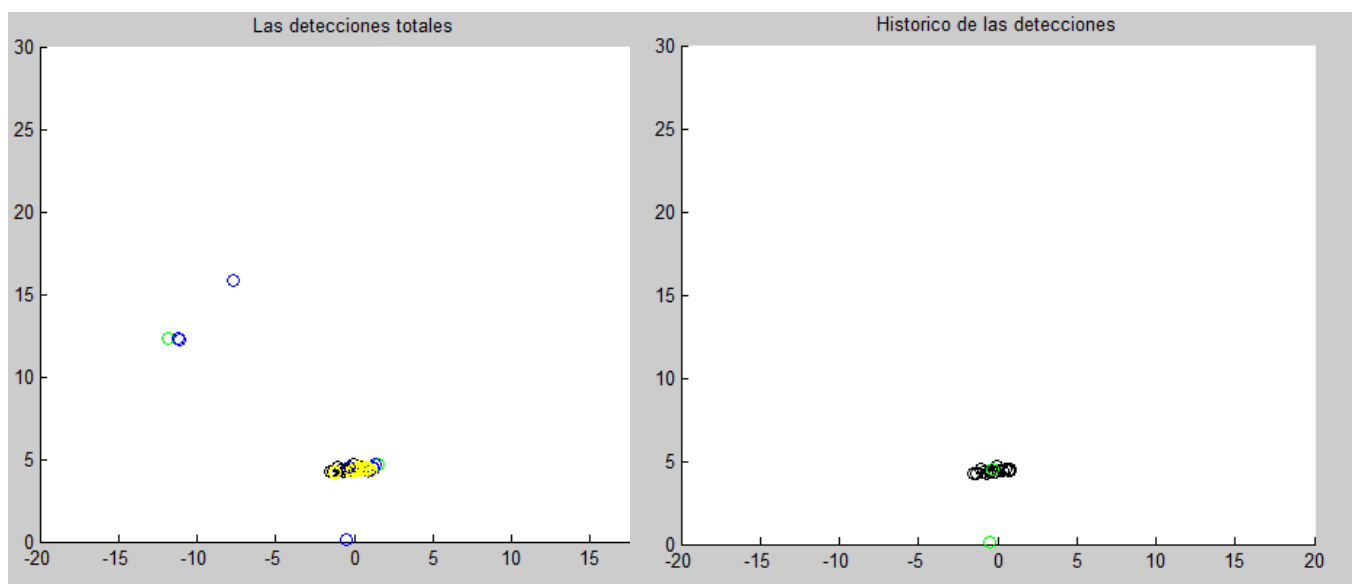
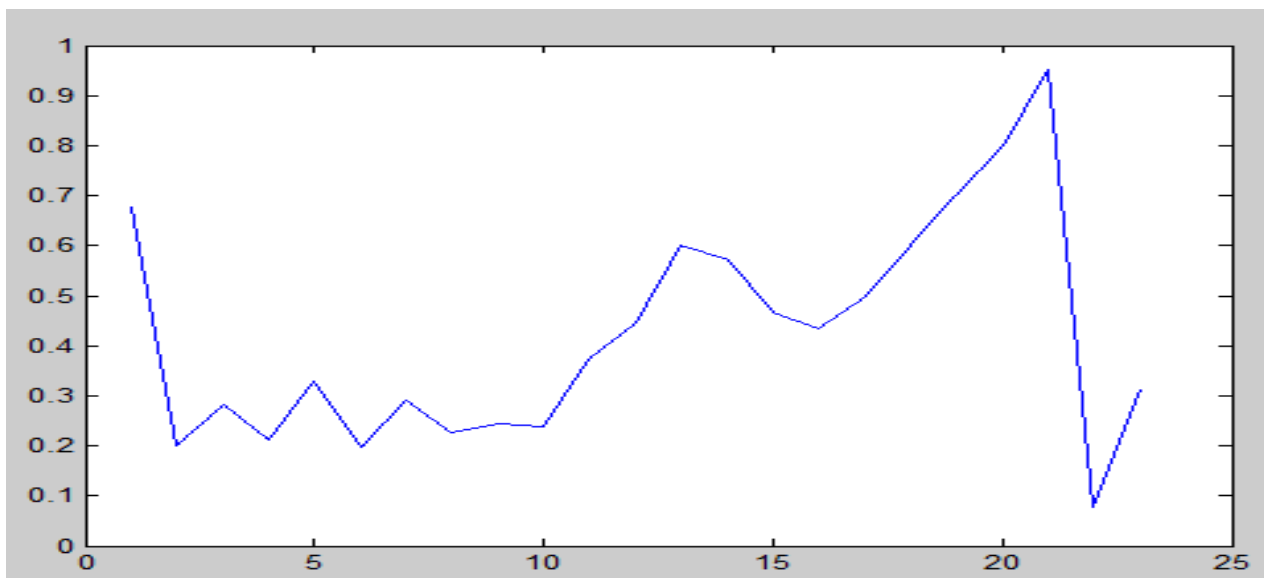
Nº de frames de la cámara (información disponible de la cámara): 14

Tabla 10. Secuencia 1PF

| | Láser | Cámara | Fusión |
|--------------------|---------------|---------------|---------------|
| Positivos | 25 | 6 | 19 |
| % Positivos | 96.15% | 42.85% | 73.07% |
| Falsos positivos | 3 | 0 | 0 |
| % Falsos positivos | 11.53% | 1.36% | 0.8% |

Error total de la secuencia: 0,377

Error por frame:



5.3.2.2 Secuencia 2

A 10 metros estos son los resultados obtenidos por el filtro de Kalman Unscented:

Nº de frames totales: 102

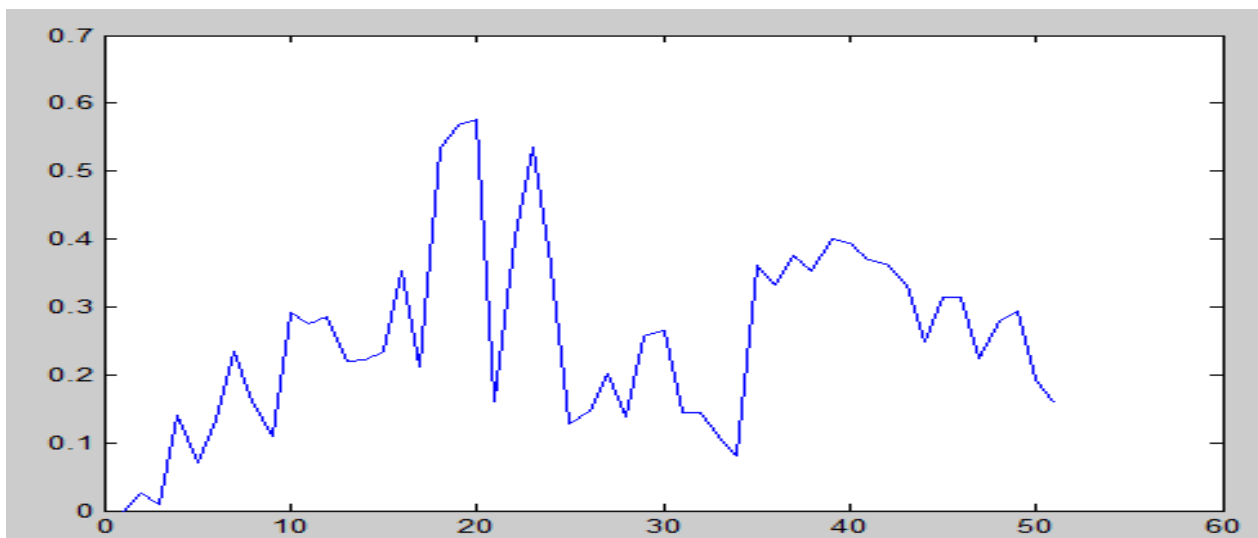
Nº de frames de la cámara (información disponible de la cámara): 54

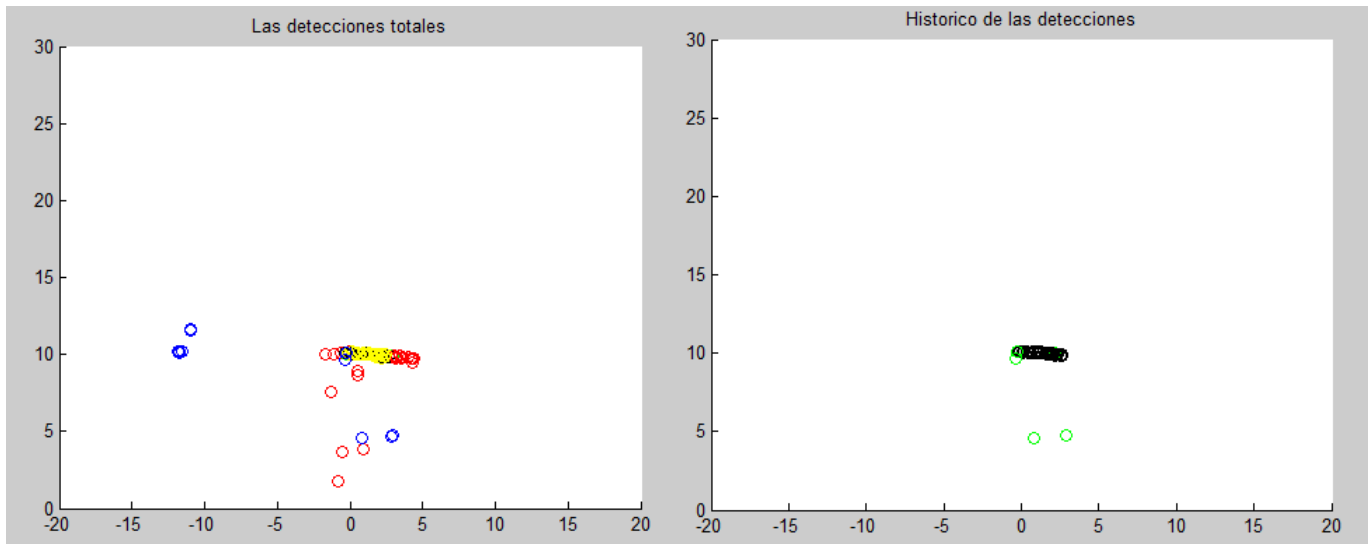
Tabla 11. Secuencia 2 PF

| | Láser | Cámara | Fusión |
|--------------------|--------|--------|--------|
| Positivos | 42 | 24 | 43 |
| % Positivos | 41.17% | 44.44% | 42.15% |
| Falsos positivos | 9 | 1 | 2 |
| % Falsos positivos | 8.82% | 1.85% | 1.96% |

Error total de la secuencia: 0,1969

Error por frame:





Como se puede observar, los resultados obtenidos por los 3 filtros son similares, por lo tanto descartamos analizar los demás peatones con el filtro de Kalman Unscented y el filtro de Partículas.

Capítulo 6. Conclusiones y trabajos futuros

6.1 Conclusiones.

En este proyecto se ha realizado un amplio y detallado estudio de las distintas técnicas de fusión sensorial para detección y seguimiento de peatones existentes actualmente y, basándonos en ellos, se ha diseñado un sistema de detección y seguimiento de peatones con un escáner láser y una cámara de visión estéreo.

Con los resultados obtenidos mediante las pruebas se ha demostrado que el sistema es bastante fiable reuniendo los dos subsistemas de detección y seguimiento de peatones. Se ha comprobado así la efectividad de los sensores utilizados y cómo esta mejora al fusionar los datos de ambos, con un porcentaje de positivos de más del 80,5%.

Por otro lado hemos conseguido bajo índice de falsos positivos en cada una de las secuencias analizadas, con un índice de 1,64% de falsos positivos.

6.2. Trabajos futuros.

Los posibles trabajos futuros serían:

Si en un futuro se le hiciera al vehículo de pruebas su correspondiente modificación tendría que ser capaz de realizar una frenada de emergencia por sí sólo si fuera necesario.

Se podría probar el sistema en entornos donde las condiciones climatológicas sean adversas como, lluvia, nieve, niebla u otras situaciones que se pueden presentar en los entornos viarios como humo y polvo denso o condiciones de baja visibilidad.

Otro posible trabajo futuro sería experimentar con el sistema incorporado en un vehículo móvil y circulando por un vía abierta o cerrada a la circulación, en las que encontremos todo tipo de situaciones (cambios de rasante, vías con firme irregular o con badenes), con un tráfico real o en su defecto lo más parecido posible. Con correcciones de la posición y de la velocidad mediante GPS u otra tecnología y un sistema de inerciales para evitar datos incorrectos cuando el automóvil atravesase terrenos irregulares y se produzca una oscilación vertical del radar láser.

Y por último se podría realizar un sistema que detectase otro tipo de objetos, ya sean otros automóviles, bicicletas y motocicletas, camiones, etc...

Capítulo 7. Presupuesto

7.1 Coste de material

En esta sección se incluye el coste de los materiales necesarios para el desarrollo del sistema de detección y seguimiento de peatones.

Tabla 12. Costes de material

| Concepto | Coste unitario | Cantidad | Total |
|-----------------------------------|----------------|----------|---------------|
| Telómetro Láser Sick LMS 291 | 3.500€ | 1 | 3.500€ |
| Cámara estéreo Bumblebee2 | 1.810€ | 1 | 1.810€ |
| GPS + Sensor Inercial Xsens MTI-G | 2.500€ | 1 | 2.500€ |
| Matlab R2007b | 550€ | 1 | 550€ |
| Portátil Acer Aspire 5735z | 530€ | 1 | 530€ |
| Total | | | 8.890€ |

7.2 Costes de personal

En esta sección se incluye el coste del personal involucrado en el desarrollo del proyecto durante el tiempo de desarrollo del mismo.

Tabla 13. Costes de personal

| Concepto | Coste unitario | Cantidad | Total |
|------------------------------|----------------|----------|---------|
| Ingeniero Técnico Industrial | 2.000€/mes | 10 meses | 20.000€ |
| Total | | | 20.000€ |

7.3. Presupuesto Total

Tabla 14. Coste total del Proyecto

| Concepto | Total |
|--------------------|---------|
| Costes de material | 8.890€ |
| Costes de personal | 20.000€ |
| Presupuesto Total | 28.890€ |

Capítulo 8. Bibliografía

Referencias:

1. http://www.dgt.es/was6/portal/contenidos/es/seguridad_vial/estadistica/publicaciones/accidentes_urbana/accidentes_trafico002.pdf (nov. 2011)
2. http://imagenes.w3.racc.es/uploads/file/12705_NP_pasos_peatones_espanya.pdf (nov. 2011)
3. U. Hofmann, A. Rieder, and E.D. Dickmanns, "Radar and vision data fusion for hybrid adaptive cruise control on highways", in Proc. Int. Conf. Comput. Vis. Syst., Vancouver, BC Canada, 2001, pp. 125-138.
4. A. Sole, O. Mano, G. Stain, H. Kumon, Y. Tamatsu, and A. Shashua, "Solid or not solid: Vision for radar target validation," in Proc. IEEE Intell. Vehicles Sym. Parma Italy, Jun 2004, pp. 819-824.
5. Ofer, A.S., Mano, O., Stein, G.P., Kumon, H., Tamatsu, Y., Shashua, A.: Solid or not solid: Vision for radar target validation. In: IEEE Intelligent Vehicles Symposium Proceedings, 819-824, 2004.
6. Shigetaka Suzuki, Pongsathorn Raksincharoensak, Ikuko Shimizu, Masao Nagai and Rolf Adomat: "Sensor Fusion-Based Pedestrian Collision Warning System with Crosswalk Detection". In: 2010 IEEE Intelligent Vehicles Symposium University of California, San Diego, CA, USA June 21-24, 2010
7. F. García, D. Olmeda, J. M. Armingol and A. de la Escalera: " Hybrid Fusion Scheme for Pedestrian Detection based on Láser Scanner and Far Infrared

- Camera.”, In: 2010 IEEE Intelligent Vehicles Symposium University of California, San Diego, CA, USA June 21-24, 2010
8. Alberto Broggi, Pietro Cerri, Stefano Ghidoni, Paolo Grisleri, and Ho Gi Jung:”A New Approach to Urban Pedestrian Detection for Automatic Braking”, In IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 10, NO. 4, DECEMBER 2009
 9. L. Leyrit, C. Chateau, C. Tournayre, and J.-T. Lapresté, “Association of AdaBoost and kernel based machine learning methods for visual pedestrian recognition,” in Proc. IEEE Intell. Vehicles Symp., Eindhoven, The Netherlands, Jun. 2008, pp. 67–72.
 10. Matthias Serfling, Otto Loehlein, Roland Schweiger and Klaus Dietmayert: “Camera and Imaging Radar Feature Level Sensorfusion for Night Vision Pedestrian Recognition” 978-1-4244-3504-3/09/\$25.00 ©2009 IEEE
 11. Vadim Frolov and Fernando Puente León: “Pedestrian detection based on maximally stable extremal regions” In: : 2010 IEEE Intelligent Vehicles Symposium University of California, San Diego, CA, USA June 21-24, 2010
 12. S. Gidel, C. Blanc, T. Chateau, P. Checchin and L. Trassoudaine: “Non-parametric Laser and Video Data Fusion: Application to Pedestrian Detection in Urban Environment” In: 12th International Conference on Information Fusion Seattle, WA, USA, July 6-9, 2009
 13. Chao-Hua Guan, Jian-Wei Gong, Yong-Dan Chen and Hui-Yan Chen:” An application of data fusion combining laser scanner and vision in real-time driving environment recognition system”, In: Proceedings of the Eighth International Conference on Machine Learning and Cybernetics, Baoding, 12-15 July 2009

14. S. Gidel, C. Blanc, T. Chateau, P. Checchin and L. Trassoudaine: "Comparison between GMM and KDE Data Fusion methods for Particle Filtering: Application to Pedestrian Detection from Laser and Video Measurements"
15. G. Gate, A. Breheret and F. Nashashibi: "Fast Pedestrian Detection in Dense Environment with a Laser Scanner and a Camera" In: September 15, 2008.
16. Sebastian Schneider, Michael Himmelsbach, Thorsten Luettel and Hans-Joachim Wuensche: "Fusing Vision and lidar – Synchronization, Correction and Occlusion Reasoning", In: 2010 IEEE Intelligent Vehicles Symposium University of California, San Diego, CA, USA June 21-24, 2010
17. <http://www.hizook.com/projects/sick-lms-291-laser-rangefinder-lidar> (nov. 2011)
18. <https://www.mysick.com/eCat.aspx?go=DataSheet&Cat=Row&At=Fa&Cult=English&ProductID=33771> (nov. 2011)
19. http://www.ptgrey.com/products/bumblebee2/bumblebee2_stereo_camera.asp (nov. 2011)
20. <http://www.xsens.com/en/general/mti-g> (nov. 2011)
21. <http://www.xenarc.com/product/index-monitor.html> (nov. 2011)
22. <http://es.wikipedia.org/wiki/MATLAB> (nov. 2011)
23. <http://es.wikipedia.org/wiki/OpenCV> (nov. 2011)
24. http://en.wikipedia.org/wiki/Kalman_filter (ene. 2012)



25. http://en.wikipedia.org/wiki/Extended_Kalman_filter (feb. 2012)
26. http://en.wikipedia.org/wiki/Kalman_filter#Unscented_Kalman_filter (feb. 2012)
27. http://en.wikipedia.org/wiki/Particle_filter (feb. 2012)